Testing White’s Floristic Impoverishment Hypothesis in the Forest-Savanna Transition Zones of Africa

Mathew Rees 1,2,*, John L. Godlee 1, David J. Harris 2, Casey M. Ryan 1 and Kyle G. Dexter 1,2

1 School of GeoSciences, University of Edinburgh, Edinburgh EH9 3FF, UK; kyle.dexter@ed.ac.uk (K.G.D.)
2 Tropical Diversity, Royal Botanic Garden Edinburgh, Edinburgh EH5 3LR, UK
* Correspondence: mathew.rees@ed.ac.uk

Abstract: In tropical Africa, forests and savannas are the two most widespread biomes and potentially represent alternative stable states with a divergent species composition. A classic, but untested, hypothesis posited by White (1983) suggests that the transition zones between forests and savannas contain a floristically impoverished assemblage, with few representatives from each biome. Further, the evolutionary dimension of diversity has received limited attention, despite its importance for understanding the biogeographic history of biomes. Here, we quantified species richness and several measures of evolutionary diversity in 1° grid cells, using c. 300,000 occurrence records of trees and shrubs combined with biome affiliation data for 3203 species. We found that assemblages in transition zones hold fewer woody species than assemblages in forest and savanna zones, as posited by White. However, transition zones hold more phylogenetic diversity than expected given their species richness, whether one considers forest and savanna assemblages separately or together. We also show that the Congo Basin forest has low levels of phylogenetic diversity, given the number of species, and highlight south-eastern African savannas as a centre of savanna woody species richness and phylogenetic diversity. Regions with high phylogenetic diversity, given the number of both forest and savanna species, were centred around the Dahomey Gap and Cameroon, mainly in transition zones. Overall, our study shows that even if floristically impoverished, transition zones hold unexpectedly high evolutionary diversity. This suggests that they are important centres of evolutionary innovation and diversification and/or serve as evolutionary crossroads, where lineages that diversified in contrasting environments coexist within a single area.

Keywords: Africa; forest; savanna; phylogenetic; diversity; transition; woody

1. Introduction

“The floras of the Guineo-Congolian and Zambezian Regions are almost mutually exclusive. There is, however, a transition zone between them, up to 500 km wide and of considerable complexity. In it, an impoverished Guineo-Congolian flora and an even more impoverished Zambezian flora interdigitate or occur in mosaic, and locally intermingle.”

Frank White (1983) [1]

In Africa, forests and savannas are the two dominant biomes, representing over 60% of the continent’s land surface [2–4]. Both biomes provide important ecosystem services, such as carbon sequestration, biodiversity maintenance, water purification, and energy sources [5,6]. The tropical rainforests of Western and Central Africa represent the second largest tropical rainforest on Earth after the Amazon basin and host an estimated 10,000 species of vascular plants, with over 50% of lianas, trees, and shrubs being endemic [7,8]. The Miombo woodlands of Southern and Eastern Africa, the largest savanna in the world, are thought to contain around 8500 species of plants, over half of which are endemic [9]. These two major vegetation units correspond to two regional centres of endemism (RCE) described by White [1]: the Guineo-Congolian RCE and the Zambezian RCE. In his work, White [1] noted that these two RCEs tend to be dominated by very
different species assemblages and hardly shared any species in common. In the northern regions of Angola and the southern Democratic Republic of Congo (DRC), where the two floras meet, patches of vegetation tend to be dominated either by Guineo-Congolian species (forming forests) or by Zambezian species (forming savannas), but intermediate vegetation types are exceedingly rare. White [1] also suggested that these regions (hereafter-called transition zones) tend to be impoverished in both forest species and savanna species. In other words, these transition zones represent a much-reduced sub-set of their respective RCE. When we use the term “transition zone”, we mean the spatial region where both forests and savanna assemblages are present and frequently found next to one another. This is in line with White’s use of the term “transition zone”. We do not mean specific local ecotones where forest and savanna communities meet. At the time, White [1] based his statements on his expert knowledge of the flora but relatively few data points. We thus consider his assertion to be a hypothesis, which remains untested with modern empirical data sets and statistical methods.

White’s hypothesis takes as a basis that forests and savannas form mutually exclusive vegetations, later referred to as alternative stable states [10,11]. Alternative stable state theory suggests that within a given environmental space, the landscape can switch from one state to another, i.e., from a forest-dominated landscape to a savanna-dominated landscape or vice-versa, but that intermediate states are scarce to absent [12]. Forests are closed canopy ecosystems, dominated by low light levels, which increases humidity and reduces grass cover and fire. This in turn promotes the growth of trees. Savannas, on the other hand, are open canopy ecosystems, with high light levels, which increases grass cover, which then increases fire and herbivory. This in turn reduces the growth of trees and promotes more grasses [13]. The intensity and regularity of fire play an important role in shaping the transitions between forests and savannas [14,15] (for the specific structural thresholds that help categorise vegetation as either forest or savanna, see Ratnam et al. [16] and Torello-Raventos et al. [17]).

Recent work [18] supports the idea that forest and savannas have mutually exclusive tree species composition, and that both states can potentially occur across large areas of Africa (but see [19]). However, previous large-scale analyses of tree species richness patterns in Africa [7,20–25] have not distinguished forest from savanna assemblages. In order to properly test White’s hypothesis of floristic impoverishment, it is necessary to analyse forest and savanna diversity patterns separately. Only then will it be possible to determine if transition zones are indeed composed of impoverished forest and savannas assemblages. In addition, little is known about the evolutionary dimension of forest and savanna diversity patterns in Africa. One might hypothesise that transition zones have reduced evolutionary diversity, as they are more distant from the core zones in which forest and savanna floras have evolved [26,27]. Alternatively, we might expect transition zones to have elevated evolutionary diversity if they represent centres of evolutionary innovation and diversification [28] or a crossroads where lineages from different parts of the plant tree of life meet and co-occur [29].

Many scientists have sought to explain the processes that generate diversity. Allopatric speciation is perhaps the most widely cited mode of speciation [30], which is unsurprising given that the idea stems directly from the work on island biogeography initiated by Darwin [31] and subsequently developed by MacArthur and Wilson [32]. This led to the hypothesis that most species in the continental tropics originated through ecological break-ups during climatic fluctuations and retreated to climatically stable and island-like refugia [33–36]. Alternatively, parapatric speciation has received some support, where divergent natural selection along ecological clines can result in speciation even in the presence of gene flow [28,37–42]. Both models are not necessarily mutually exclusive, and the current composition of communities is likely a result of both processes (reviewed in Mortiz et al. [43]). A key concept that unites these theories is that of ecological stability vs. instability. Stebbins [44] suggested that climatically stable areas represent areas of high diversity where evolutionary lineages have accumulated over time, whilst areas of high
ecological instability represent areas of evolutionary diversification, where lineages are produced and die off quickly, resulting in high turnover rates [44,45].

The most recent understanding of angiosperm evolutionary diversification across the African continent comes from the study of Dagalier et al. [46], who combined taxonomically verified occurrence data [47] with a large phylogeny [48] to look at genus-level patterns of phylogenetic diversity and endemism in Africa. They found that the forests of Guinea-Congolia had high levels of relative phylogenetic diversity (RPD) [49], suggesting these areas are dominated by genera on long terminal branches, or evolutionarily old genera. Conversely, they found that Eastern Africa and the Dahomey Gap had lower levels of RPD, suggesting that these areas are dominated by genera on short terminal branches, or evolutionarily young genera. However, no study to date has distinguished forest assemblages from savanna assemblages when studying evolutionary diversity patterns across Africa. Doing so would allow us to test whether White’s impoverishment hypothesis holds over evolutionary timescales.

In order to address these knowledge gaps, we combined a large dataset of woody species occurrences with biome affiliation data and a phylogenetic tree to investigate taxonomic and evolutionary diversity patterns in 1° grid cells throughout tropical Africa. Specifically, we asked the following questions:

- Do transition zones contain fewer species than forest or savanna core zones?
- Do transition zones hold less phylogenetic diversity than forest or savanna core zones?
- Which regions of Africa hold the highest woody plant species richness and phylogenetic diversity for forest and savanna assemblages?

2. Methods
2.1. Species Biome

We used species biome affiliation from Gorel et al. [50] to assign each species to either the “Forest” or “Savanna” biome. Briefly, the dataset comprises 4142 species of vascular plants that were classified based on climatic variables. Gorel et al. [50] used a hierarchical clustering approach to split species into broad climatic groups. We also checked that these biome assignments driven by climate data matched up with those of Aleman et al. [18], which was based on the physiognomic vegetation types in which species predominately occurred (forest versus savanna), independent of climate data. For tree species present in both datasets, both methods gave the same biome assignment over 90% of the time (Supplementary File S1).

2.2. Species Occurrence

The RAINBIO [47] and FLOTROP [51] databases were combined to produce a clean and taxonomically verified list of georeferenced herbarium specimens. All *Acacia* names were assigned to *Senegalia* or *Vachellia* based on Kyalangalilwa et al. [52]. We filtered the data to include only the species present in the biome affiliation data from Gorel et al. [50] and selected species from the woody plant categories of tree, shrub, and shrublet (i.e., small shrub), as defined by the RAINBIO database. We removed all gymnosperms and tree ferns, as these are not well represented in the biome dataset from Gorel et al. [50] and would add substantial branch length to very few assemblages [53]. Although herbarium specimens can provide an excellent snapshot of biodiversity in a spatial context, we acknowledge that these datasets come with inherent biases [54].

2.3. Floristic Assemblages

To compare spatially delimited “assemblages” of trees and shrubs, we used 1° grid cells (~100 × 100 km at the equator) from Marshall et al. [55]. The data comprised 1197 grid cells ranging from the Tropic of Capricorn to the Tropic of Cancer (24° N, 24° S). We excluded all records located outside these grid cells. After cleaning, our main dataset contained around 300,000 records, representing 3203 species (Table S1). For each grid cell, we counted the total number of species and the proportion of forest and savanna species (Pf,
We then defined a biome threshold ($B_t$) for the $P_f$ and $P_s$, above which we considered the cell to be in the core of that zone (i.e., either core forest zone if $P_f > B_t$ or core savanna zone if $P_s > B_t$). We initially set $B_t$ to 70% but varied this from 60% to 80% to check the consistency of our results. For cells where the max ($P_f$, $P_s$) was lower than $B_t$, we classified them as a ‘transition zone’. We plotted our classification on the map of Africa to inspect if our transition zones broadly matched with known areas of transition [1,56] (Figures S5–S7).

2.4. Phylogenetic Reconstruction

We used the R package V.PhyloMaker [57] to build a phylogenetic tree from the species recorded within our study region. About 36% of our species were already present in the supertree included in V.PhyloMaker, which is based on Smith and Brown [58] with some modifications [57]. We used scenario three of V.PhyloMaker to bind our new tips to the tree. This consists of binding species that were not present in the original phylogeny to the half-way point of the family or genus branch to which they belong. Previous studies have shown that using large scale phylogenies derived from few markers and with multiple polytomies generally does not affect the conclusions drawn from large-scale community phylogenetic analyses [59,60].

2.5. Diversity Metrics

For each grid cell, we calculated species richness (SR), Faith’s phylogenetic diversity (PD), mean nearest taxon distance (MNTD), and mean pairwise distance (MPD). To account for differences in species richness among grid cells, we also calculated their standardised versions ($PD_{ses}$, $MNTD_{ses}$, $MPD_{ses}$) using the R package PhyloMeasures [61]. The standardised versions represent a measure of the excess or deficit of phylogenetic diversity from that expected given the species richness of an assemblage. We used the “uniform” null model repeated 1000 times. This model draws N species at random equal to the SR of the observed community. The mean expected value from the null distribution is then subtracted from the observed value and divided by the standard deviation of the mean from the null distribution. For each grid cell, we did this for all species combined and for forest and savanna species separately; thus, the species pool considered was either the whole dataset (3203 species), all forest species (1883 species), or all savanna species (1320 species), respectively. PD is a measure of diversity linked to the “richness” dimension of phylogenetic diversity, whilst MNTD and MPD are linked to the “divergence” dimension of phylogenetic diversity [62]. PD and MNTD tend to be correlated because they are both sensitive to the “shallow” nodes of the phylogenetic tree, whilst MPD measures divergence at deeper nodes in the tree [63].

2.6. Statistical Analyses

As the sampling effort and number of species varied greatly between grid cells, prior to the calculation of diversity metrics, we rarefied our grid cells by number of collections, using 30 collections as the minimum threshold to include a grid cell. We also varied this from 20 to 40 collections to check the consistency of our results. This method preserves the species richness gradient between grid cells but controls for the number of occurrences included in each grid cell (as species richness and collection number were highly correlated). We repeated the rarefactions 100 times and calculated the mean of each metric across rarefactions. From here on, when we refer to these diversity metrics, we mean the rarefied metrics.

We used violin plots to visually compare diversity metrics between our three zones (forest vs. transition vs. savanna). To test whether diversity metrics varied among zones, we first tested for homogeneity of variance using Levene’s tests. To control for non-normally distributed variance between our groups, we used the Kruskal-Wallis test and post-hoc Dunn’s test using the Holm correction for multiple pairwise comparisons. Standardized effect sizes (SES) were computed using Mangiafico’s d in the R package rcompanion [64]. Mangiafico’s d is analogous to Cohen’s d but it uses the difference in medians divided by
the pooled median absolute deviation (MAD). We used the 10th and 90th percent quantiles to evaluate the lowest and highest areas of diversity, respectively, for each metric.

All data manipulations and analyses were performed in the R statistical Environment v.4.1.1 [65].

3. Results

3.1. Species Richness

Forest assemblages showed the highest rarefied species richness (SR) in core forest zones \( (n = 203, \text{median} = 26.86, \text{MAD} = 2.02) \), followed by transition zones \( (n = 68, \text{median} = 22.59, \text{MAD} = 3.08) \), and then core savanna zones \( (n = 57, \text{median} = 15.64, \text{MAD} = 5.9) \) (Table S2). The Kruskal-Wallis test showed a significant difference between the medians of each zone \( (\chi^2 = 177.86, \text{df} = 2, p < 0.01) \), and the post-hoc Dunn test showed a significant difference between each pair \( (p < 0.01 \text{ between all three pairs, forest vs. transition SES = 1.64, forest vs. savanna SES = 2.54, transition vs. savanna SES = 1.48}) \) (Figure 1A). Geographically, forest assemblage SR was highest in Upper Guinea, specifically in the countries of Liberia, Ivory Coast, and Ghana; and in Lower Guinea and Congolia, specifically in Gabon, Equatorial Guinea, Cameroon, southern Nigeria, the Republic of Congo (ROC) and the Democratic Republic of Congo (DRC) (Figure 2A).

Figure 1. Violin plots of rarefied species richness (SR) for forest assemblages (A) and savanna assemblages (B), rarefied phylogenetic diversity (PD) for forest assemblages (C) and savanna assemblages (D), and rarefied standardised effect size of phylogenetic diversity (PD ses) for forest assemblages (E) and savanna assemblages (F) across core forest, transition, and core savanna zones.

The significant pairwise comparison of medians using the Dunn test with Holm correction are plotted as bars with p-values.
assemblages (D), and rarefied standardised effect size of phylogenetic diversity (PDses) for forest assemblages (E) and savanna assemblages (F) across core forest, transition, and core savanna zones. The significant pairwise comparison of medians using the Dunn test with Holm correction are plotted as bars with $p$-values.

**Figure 2.** Maps of rarefied species richness (SR) for forest assemblages (A) and savanna assemblages (B), rarefied phylogenetic diversity (PD) for forest assemblages (C) and savanna assemblages (D), and rarefied standardised effect size of phylogenetic diversity (PDses) for forest assemblages (E) and savanna assemblages (F).

Savanna assemblages showed the highest rarefied SR in core savanna zones ($n = 424$, median = 23.42, MAD = 4.92), followed by transition zones ($n = 71$, median = 22.26, MAD = 2.64), and then core forest zones ($n = 64$, median = 18.93, MAD = 3.74) (Table S2). The Kruskal-Wallis test showed a significant difference between the medians of each zone...
(\chi^2 = 29.2, df = 2, p < 0.01), and the post-hoc Dunn test showed that core savanna and transition zones had a significantly higher SR compared to core forest zones (p < 0.01, savanna vs. forest SES = 1.03, transition vs. forest SES = 1.03). The median SR of transition zones was lower than that of core savanna zones, but this difference was not significant (p = 0.2) (Figure 1B). Geographically, the savanna assemblage SR was highest in the southeast part of our study area, in Tanzania and Mozambique (Figure 2B).

3.2. Phylogenetic Diversity

Forest assemblages showed the same PD patterns as SR (Figure 1B). Meanwhile, the highest PDses for forest assemblages was in transition zones (n = 68, median = 0.16, MAD = 0.75), followed by core savanna zones (n = 57, median = -0.02, MAD = 0.9), and then core forest zones (n = 203, median = -0.45, MAD = 1.13) (Figure 1E, Table S2). The Kruskal-Wallis test showed a significant difference between the medians of each zone (\chi^2 = 27.9, df = 2, p < 0.01), and the post-hoc Dunn test showed that core savanna and transition zones had significantly greater PDses compared to core forest zones (p < 0.01, savanna vs. forest SES = 0.42, transition vs. forest SES = 0.63 respectively). Transition zone PDses was not significantly different from that of core savanna zones (p = 0.4) (Figure 1E). MPDses for forest assemblages was highest in transition zones (n = 68, median = 0.07, MAD = 0.74) and lowest in core savanna zones (n = 57, median = -0.36, MAD = 0.7), with transition zones showing significantly higher median MPDses from core forest and core savanna zones (p < 0.01, transition vs. forest SES = 0.37, transition vs. savanna SES = 0.6). Core forest and core savanna zones were not significantly different (p = 0.39) (Figure S1). MNTDses for forest assemblages was highest in core savanna zones (n = 57, median = 0.24, MAD = 0.88), followed by transition zones (n = 68, median = 0.09, MAD = 0.72), and then core forest zones (n = 203, median = -0.47, MAD = 0.88). Core savanna and transition zones MNTDses were not significantly different from each other (p = 0.39), but both were greater than core forest zones (p < 0.01, savanna vs. forest SES = 0.81, transition vs. forest SES = 0.7) (Figure S1, Table S2).

Geographically, forest assemblage PD followed similar patterns to SR (Figure 2C). Meanwhile, PDses was the highest in six areas: along the coast of Guinea-Bissau; in Ivory-Coast and the Dahomey Gap; in Eastern Cameroon; near the Reserve de Chasse de la Lefini in the ROC; between the Massif d’Itombwe and Mahale Mountain in the southern part of the Albertine rift; and along the northern limits of the Eastern Arc mountains and the coastal forests of Tanzania and southern Kenya. Forest assemblage PDses was lowest in the central Congo Basin (Figures 2E, 3 and S8).

Savanna assemblages showed the highest PD in transition zones (n = 71, median = 1732, MAD = 150), followed by core savanna zones (n = 424, median = 1639, MAD = 433), and finally core forest zones (n = 64, median = 1459, MAD = 257) (Figure 2D). Core forest zones had significantly lower PD than transition and core savanna zones (p < 0.01, forest vs. transition SES = 1.29, forest vs. savanna SES = 0.5), but transition vs. core savanna zones were not significantly different from each other (p = 0.4). The PDses of savanna assemblages was highest in transition zones (n = 71, median = -0.01, MAD = 0.95), followed by core forest zones (n = 64, median = -0.49, MAD = 0.97), and then core savanna zones (n = 424, median = -0.61, MAD = 0.95) (Figure 1F, Table S2). The Kruskal-Wallis test showed a significant difference between the medians of each zone (\chi^2 = 29, df = 2, p < 0.01). The transition zone PDses was significantly higher than either core forest or core savanna zones (p < 0.01, transition vs. forest SES = 0.5, transition vs. savanna SES = 0.63). Core forest and core savanna zones were not significantly different from each other (p = 0.4) (Figure 1F). The MPDses of savanna assemblages was highest in transition zones (n = 71, median = 0.18, MAD = 0.83), followed by core forest zones (n = 64, median = -0.4, MAD = 1.06), and then core savanna zones (n = 424, median = -0.49, MAD = 0.98), with transition zones showing significantly higher medians from core forest and core savanna zones (p < 0.01, transition vs. forest SES = 0.61, transition vs. savanna SES = 0.75). Core forest and core savanna zones were not significantly different (p = 0.35) (Figure S1). The MNTDses of savanna assemblages were not significantly different from each other (p = 0.2) (Figure 1E), with transition zones showing significantly higher median MNTDses from core forest and core savanna zones (p < 0.01, transition vs. forest SES = 0.37, transition vs. savanna SES = 0.6). Core forest and core savanna zones were not significantly different (p = 0.39) (Figure S1). The MNTDses of savanna assemblages were not significantly different from each other (p = 0.4) (Figure 1F). The MPDses of savanna assemblages was highest in transition zones (n = 71, median = 0.18, MAD = 0.83), followed by core forest zones (n = 64, median = -0.4, MAD = 1.06), and then core savanna zones (n = 424, median = -0.49, MAD = 0.98), with transition zones showing significantly higher medians from core forest and core savanna zones (p < 0.01, transition vs. forest SES = 0.61, transition vs. savanna SES = 0.75). Core forest and core savanna zones were not significantly different (p = 0.35) (Figure S1). The MNTDses of savanna assemblages were not significantly different from each other (p = 0.2) (Figure 1E), with transition zones showing significantly higher median MNTDses from core forest and core savanna zones (p < 0.01, transition vs. forest SES = 0.37, transition vs. savanna SES = 0.6). Core forest and core savanna zones were not significantly different (p = 0.39) (Figure S1). The MNTDses of savanna assemblages were not significantly different from each other (p = 0.4) (Figure 1F). The MPDses of savanna assemblages was highest in transition zones (n = 71, median = 0.18, MAD = 0.83), followed by core forest zones (n = 64, median = -0.4, MAD = 1.06), and then core savanna zones (n = 424, median = -0.49, MAD = 0.98), with transition zones showing significantly higher medians from core forest and core savanna zones (p < 0.01, transition vs. forest SES = 0.61, transition vs. savanna SES = 0.75). Core forest and core savanna zones were not significantly different (p = 0.35) (Figure S1). The MNTDses of savanna assemblages were not significantly different from each other (p = 0.2) (Figure 1E), with transition zones showing significantly higher median MNTDses from core forest and core savanna zones (p < 0.01, transition vs. forest SES = 0.37, transition vs. savanna SES = 0.6). Core forest and core savanna zones were not significantly different (p = 0.39) (Figure S1). The MNTDses of savanna assemblages were not significantly different from each other (p = 0.4) (Figure 1F). The MPDses of savanna assemblages was highest in transition zones (n = 71, median = 0.18, MAD = 0.83), followed by core forest zones (n = 64, median = -0.4, MAD = 1.06), and then core savanna zones (n = 424, median = -0.49, MAD = 0.98), with transition zones showing significantly higher medians from core forest and core savanna zones (p < 0.01, transition vs. forest SES = 0.61, transition vs. savanna SES = 0.75). Core forest and core savanna zones were not significantly different (p = 0.35) (Figure S1).
was highest in transition zones \((n = 71, \text{median} = -0.05, \text{MAD} = 0.76)\), followed by core forest zones \((n = 64, \text{median} = -0.52, \text{MAD} = 0.74)\), and then core savanna zones \((n = 424, \text{median} = -0.66, \text{MAD} = 0.84)\). Core savanna and core forest zone MNTDses values were marginally different \((p = 0.06)\), but both were lower compared to transition zones \((p < 0.01, \text{savanna vs. transition SES} = 0.76, \text{forest vs. transition SES} = 0.63)\) (Figure S1, Table S2).

**Figure 3.** Map of highest 10% standardised effect size of phylogenetic diversity (PDses) for forest assemblages (green) and savanna assemblages (brown). Blue outlines represent transition zones.

Geographically, the pattern of savanna assemblages’ PD was similar to that of SR (Figure 2D). The PDses of savanna assemblages was highest in four areas: along the Dahomey Gap; along the volcanic line in Cameroon and east of Lake Mbakaou; in the Ethiopian highlands; and finally along coastal Tanzania and Mozambique. The lowest savanna assemblage PDses did not follow a clear geographic pattern but could be found scattered across Eastern and Southern Africa (Figures 2F, 3 and S9).

4. Discussion

In this study, we combined large datasets of taxonomically verified species occurrence from herbarium specimens with biome affiliation data and a phylogenetic tree to distinguish and map diversity patterns for forest and savanna tree and shrub assemblages across tropical Africa. The results showed a decoupling of species richness and phylogenetic diversity in forest–savanna transition zones. We found that forest and savanna assemblages showed the highest rarefied species richness in their respective core zones of occurrence (i.e., areas dominated by either forest or savanna species), confirming White’s hypothesis. However, the highest phylogenetic diversity given the number of species for both forest and savanna assemblages was in transition zones, suggesting they are important eco-evolutionary laboratories for woody plant diversity and/or represent evolutionary crossroads, where lineages that diversified in contrasting environments coexist within a single area. Finally, the highest excess of phylogenetic diversity given the species richness for both forest and savanna assemblages was concentrated around the Dahomey Gap and Cameroon, mainly within transition zones.
4.1. Forest Diversity

Our results showed a decline in forest SR as we moved from core forest zones to transition zones and then core savanna zones, thus supporting White’s hypothesis of floristic impoverishment in transition zones (Figure 1A). Grid cells with the highest SR were concentrated in Upper Guinea, Lower Guinea, and Congolia (Figures 2A, S3 and S8), a result reflected in previous studies [7,8,22,23,27,55]. Most of those studies (except [27]) do not separate patterns for forest and savanna species, yet the overall spatial pattern of species richness reflects what we found for forest species here, presumably because forest species numerically dominate the overall pool of tree and shrub species in Africa [66].

Forest biogeography has received considerable attention, with several studies highlighting the traces of climatic fluctuation and disturbance throughout the late Paleogene and Neogene in Guinea and Congolia [67,68]. Forests expanded and contracted repeatedly during this period. This is partly due to the glacial-interglacial Milankovitch cycles, with periods of cool and dry atmosphere pushing forests back towards areas with sufficient water [69]. These events have left traces in both the fossil record and the molecular record [69–73]. For example, fossil pollen around the Turkana basin in East Africa, which is currently composed of open vegetation, indicates an increase in taxa associated with closed canopy forests around 14,000 years ago, which coincides with the onset of the African Humid Period [74,75]. Couvreur et al. [70] used dated molecular phylogenies to highlight the multiple break-ups of the Congo Basin forests. With the advent of molecular biology and the availability of large phylogenetic trees to conduct macroecological studies, we were able to push White’s hypothesis one step further and look into the phylogenetic diversity of assemblages across Africa.

Core forest zones showed the lowest median PDses values (Figure 1E), suggesting that, on average, they tend to have less phylogenetic diversity than expected, given their species richness, and tend to be more phylogenetically clustered. Our analyses of MPDses and MNTDses support this view, with forest zones showing clustering both at shallow and deeper nodes of the phylogenetic tree (Figure S1). This means that forests zones tend to have more closely related forest lineages and species and that these species tend to be younger and found on shorter branches.

Contrasting with the results of Dagalier et al. [46], who found that, at the genus level, the Congo Basin had high levels of relative phylogenetic diversity (RPD), our results showed that, at the species level, the Congo Basin had the lowest levels of phylogenetic diversity given the number of forest species. This could suggest that many of the genera in the Congo Basin are old, but that the species that compose them are young and often belong to the same clades (clustering). This is reflected by the high rate of endemism of tree species in the Congo Basin, with over 30% of forest tree species being found nowhere else in the world [7]. Together, this supports the evidence that the Congo Basin has experienced multiple contractions and expansions, both in recent and deep past [67,69,76].

Forest assemblages in savanna zones, however, tended to show low MPDses values, meaning clustering at the deeper nodes, but high MNTDses values, meaning over-dispersion at the shallow nodes of the phylogenetic tree. This suggests in turn that savanna zones tend to share more closely related forest lineages than core forest or transition zones, but that those lineages might have older species, sustained by longer branches, which would increase their relative distance to their most closely related forest species in the assemblage. This could represent the effects of environmental filtering at deep evolutionary time scales and competition between close relatives at shallow evolutionary time scales [63,77].

It is worth noting that, although most of East Africa falls within our core savanna zone, some areas are well known to support highly diverse and ancient forest ecosystems [3,78,79]. In fact, East Africa supports some of the highest PDses values observed, particularly in the northern limits of the Eastern Arc mountains and the coastal forests of Tanzania and southern Kenya (Figures 2E and S8). These regions have long been thought of as relictual elements of an once Pan-African forest that spread the entire width of the continent [69,80,81]. The long-term climatic stability of this region is thought to best explain
the patterns of richness and endemism of these floristic refugia [78,82]. This is partly due to the proximity to the Indian Ocean that has provided a relatively high and constant moisture throughout time [83]. Lovett [78] also suggested that environmental heterogeneity due to complex geology was important in allowing the movement of species in and out of optimal habitats when climatic changes modified the landscape.

4.2. Savanna Diversity

For savanna assemblages, our results also confirmed White’s hypothesis, where savanna species richness tends to decline as we shifted from core savanna zones to transition zones and finally core forest zones (Figure 1B). Savanna assemblages in core savanna zones tended to show more phylogenetic clustering, both at the shallow and deep nodes of the phylogeny, with low MNTD_{ses} and MPD_{ses} values (Figure S1), suggesting that they are composed of closely related lineages and species. Parallel to the pattern for forest species, this suggests diversification of savanna species within the core savanna zones.

Grid cells with the highest SR were concentrated in Tanzania and Mozambique Figures 2B, S4 and S9). The striking difference in species richness between south-eastern and north-western savannas is reflected in Linder [27], who used a clustering approach based on the shared geographical distribution of plant genera to delimit richness patterns in Africa, divided into lowland forest, savanna, tropic montane, and austro-temperate floras. These differences between south-east and north-west African savannas are also reflected in their species composition. Fayolle et al. [84] used an extensive dataset from plot data to analyse patterns of beta diversity in savannas. They found a clear separation between western and eastern African savannas, both at the species and genus level. Together with the fossil records of past vegetation in East Africa [76], our PD results seem to paint a picture of East Africa, and particularly Tanzania and Mozambique, as a centre of speciation and diversification for savannas. Indeed, recent studies on the genus Brachystegia Benth., an emblematic taxon of African savannas, found that the Miombo clade likely originated in Eastern Africa [85,86].

In contrast to the low levels of SR in western savannas, our results from PD_{ses} seem to indicate that western savannas have high levels of phylogenetic diversity, given their species richness. For instance, PD_{ses} values in the Dahomey Gap and in Cameroon along the volcanic line and east of Lake Mbakaou were particularly high. These results might suggest that West Africa could have independently evolved multiple savanna lineages, but that these lineages did not speciate or spread much outside of West Africa. Another possibility could be that extinction rates in West Africa were higher. West Africa is less topographically diverse than East Africa; thus, in the event of a climatic disturbance, there are less possibilities for assemblages to find adequate refugia. In order to investigate the extent of shared origins of savanna assemblages between East and West Africa, analyses of phylogenetic beta diversity should be explored.

4.3. Transition Zones as Centres of Evolutionary Diversity

Our results show that transition zones tend to have more phylogenetic diversity than expected, given their species richness in both forest and savanna assemblages. When both forest and savanna lineages are considered together, this can be viewed as a merging effect of phylogenetic diversity, or as Neves et al. [29] coined it, the “Environmental Crossroads Hypothesis”. This is where lineages that are adapted to opposite extremes of an environmental spectrum coexist; thus, the landscape will support a higher number of evolutionary lineages. For example, in the Neotropics, Neves et al. [29] and Griffiths et al. [87] found that phylogenetic diversity peaks at intermediate precipitation and elevation, respectively. In Africa, a recent study by Rejou-Mechain et al. [88] found that transitional forests had higher levels of phylogenetic diversity compared to central parts of the Congo Basin.

When forests and savannas are considered separately, transition zones also display higher levels of PD_{ses} than either core forest or core savanna zones. This might be best explained by higher MNTD_{ses} and MPD_{ses} in transition zones for both forest and savanna
assemblages, which suggests that species that inhabit these regions are separated by longer branches on the phylogeny and tend to come from different lineages. Transition zones are highly dynamic regions, where the patterns of vegetation can change rapidly, offering colonisation opportunities.

The adjacency of biomes in zones of transition and the frequency of change in the mosaics of vegetation over long evolutionary timescales are also potential drivers of evolutionary biome shifts [89]. The repeated contraction and expansion of forests and savannas in transition zones, following the cooling and warming of the Earth, should provide more frequent opportunities for plant lineages to speciate and produce novel features. Several studies of birds [39,90], lizards [41,42], and butterflies [91] have highlighted the role of ecotones in generating morphological, genetic, and physiological divergence through natural selection (although, see Coyne and Orr 2004, pp.118–124). Unfortunately, we still lack such studies in plants, but we predict that similar processes may act on shaping plant assemblages. As the results from Gorel et al. [50] show, biome shifts are ubiquitous across African angiosperms, suggesting that the genetic changes required to make these shifts are accessible to many lineages. These shifts tend to be more frequent from forest to savanna, a pattern that is particularly evident in one of the dominant lineages of African savannas today, the Detarioid Legumes [92]. Thus, for savanna assemblages, we might expect that transition zones are regions of recruitment of phylogenetic innovation from forest lineages. For forest assemblages in transition zones, we can hypothesize that disturbance from cyclical climatic warming and cooling presented opportunities for lineages to colonize new territories and move into new ecological niches. Thus, transition zones would hold higher levels of phylogenetic diversity.

In our analyses, two regions of transition stand out as supporting high levels of PDses for both forest and savanna assemblages: in Cameroon, along the volcanic line and east of Lake Mbakaou, and in the Dahomey Gap. It is well recognised that the volcanic line between Nigeria and Cameroon in Central Africa represents a biodiversity hotspot [93], which corresponds to a glacial refuge for forest species [67,73,94]. However, our results show that this is also an important area for savanna assemblages with high levels of phylogenetic diversity given the number of species. This region is a mosaic of forests and savannas, supporting deep gallery forests in topographical depressions, whilst savannas dominate the higher grounds. Another important region of high PDses for both forest and savanna assemblages is the Dahomey Gap (Figures 3, S8 and S9). A recent study from Adjossou et al. [95] suggested that the mountain range that separates Togo and Ghana has served as a minor glacial refuge, sharing a high proportion of its species with other refugia. The dense reticulate network of riparian forest, along with the high topographic and geomorphological heterogeneity, would have allowed species to move in and out of optimal habitats during periods of climatic change. This might suggest that the Dahomey Gap, and the northern Cameroonian belt, with their mosaics of vegetation have served as sites of diversification for savanna species when drier periods prevailed and for forest species when wetter climates united the upper and lower Guinean forest floras with the greater Congolian Basin.

In this context, we suggest that forest–savanna transition zones in Africa represent important eco-evolutionary laboratories for woody plant diversity, where lineages from across the plant tree of life have independently evolved novel traits for colonising new niches. These areas should receive more attention, as they are important for the generation and maintenance of biodiversity [96,97].

4.4. Caveats

It is important to point out that our study only looked at patterns of diversity in woody species and did not take into account the herbaceous stratum. It is likely that dynamics that operate on woody species are different from those that operate on herbaceous vegetation. When further data become available for categorising the biome preference of different herbaceous species, it will be important to re-evaluate our results. Although our analyses
used the largest number of species to date to investigate differences between forests and savannas in tropical Africa, with 3203 woody species used in this study, we only looked at a fraction of the c. 45,000 plant species (including woody and herbaceous) in tropical Africa [27]. This highlights the need for additional studies to help better understand the ecological and evolutionary dynamics of more species across the continent.

Another issue pertains to the sampling coverage. As we used occurrence data from herbarium specimens, some areas were collected from far more than others [98]. Species richness was highly correlated with sampling intensity. Although we attempted to deal with this issue by rarefying our data, the distribution of species in each grid cell was left skewed, potentially reflecting the bias of botanical collectors towards rare species. This pattern did not disappear as we increased the threshold for rarefaction. Thus, the results obtained from rarefying assemblages down to 100 samples were broadly similar to the results obtained from rarefying down to 30 individuals. The benefit of rarefying down to 30 individuals was that we were able to include more sampling units, covering a broader geographical and environmental range. A future avenue for exploring diversity patterns would be to use sample plot data to obtain a less-biased estimate of species diversity in any given region.

5. Conclusions

Although patterns of plant diversity in Africa are well studied, there remain many avenues to investigate the processes that shape plant assemblages. Here, we provide the first continental-wide study of the forest–savanna transition zones from an ecological and phylogenetic perspective. Our results support Frank White’s hypothesis of floristic impoverishment in transition zones from a species richness perspective. We extended White’s hypothesis to an evolutionary dimension and found that transition zones show elevated phylogenetic diversity given their species richness, a result that confirms that these regions are important for the generation and maintenance of biodiversity. Further work on the mechanisms that drive these patterns is needed. We suggest that studies of permanent sample plots and phylogenetic beta diversity between forests and savannas will provide further insights into the processes shaping African plant biodiversity.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/d15070833/s1, Supplementary File S1. Ecological niche check; Table S1: Cleaned data. Table S2: Results from the diversity analyses showing median values of grid cells rarefied to 30 occurrences. Figure S1: Violin plots of rarefied standardized effect size of mean pairwise distance (MPDses) for forest species (A), savanna species (B) and rarefied standardized effect size of mean nearest taxon distance (MNTDses) for forest species (C), savanna species (D) across core forest, transition and core savanna zones. Significant pairwise comparison of medians using Dunn test with Holm correction are plotted as bars with p.values. Figure S2: Map of non-rarefied woody species richness in Africa based on the 3203 species included in this study. Figure S3: Map of non-rarefied forest woody species richness in Africa. Figure S4: Map of non-rarefied savanna woody species richness in Africa. Figure S5: Map of Africa showing the classification of core forest (green) vs. transition (blue) vs. core savanna zones (brown). Figure S6: Map showing White’s transition zones (in red) vs transition zones in this paper (in blue) (adapted from White, 1983). Figure S7: Map showing forest-savanna mosaics from Ecoregions vs transition zones in this paper (adapted from Dinnerstein et al., 2017). Figure S8: Map of highest 10% SR (A), highest 10% PDses (B) and lowest 10% PDses (C) for forest assemblages. Figure S9: Map of highest 10% SR (A), highest 10% PDses (B) and lowest 10% PDses (C) for savanna assemblages
Author Contributions: M.R.: conceptualisation, methodology, formal analysis, investigation, data curation, writing—original draft, visualisation. J.L.G.: writing—review and editing, supervision. D.J.H.: writing—review and editing, supervision. C.M.R.: writing—review and editing, supervision. K.G.D.: conceptualisation, methodology, writing—review and editing, supervision. All authors have read and agreed to the published version of the manuscript.

Funding: Funding for this research was provided by NERC through an E4 DTP studentship (NE/S007407/1).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

Diversity 2023, 15, 833


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.