

# Distinct edaphic habitats are occupied by discrete legume assemblages with unique indicator species in the Cape Peninsula of South Africa

*Meshack N. Dlodlu\**, *Samson B. M. Chimphango*,  
*Charles H. Stirton* and *A. Muthama Muasya*

*Department of Biological Sciences, University of Cape Town, Private Bag X3, Rondebosch 7701, Cape Town, South Africa*

\*Correspondence address. Department of Biological Sciences, University of Cape Town, Private Bag X3, Rondebosch 7701, Cape Town, South Africa. Tel: +27-21-650-3684, Fax: +27-21-650-2301; E-mail: [dldmes001@myuct.ac.za](mailto:dldmes001@myuct.ac.za)

## Abstract

### **Aims**

The Cape Peninsula is a small area (471 km<sup>2</sup>) situated on the south-westernmost tip of the Core Cape Subregion (CCR) of South Africa. Within the Cape Peninsula, Fabaceae are the third most species-rich plant family (162 species) and they have the second highest number of endemic species after the Ericaceae. However, legumes are not the dominant taxa in the vegetation. They tend to show patchy distributions within the landscape and different species assemblages usually occupy particular niches at any given locality. The present study undertook to establish if edaphic factors influence legume species distribution in the Cape Peninsula and to determine the key indicator species for the different assemblages.

### **Methods**

Soils from 27 legume sites, spanning all major geological substrates of the Cape Peninsula, were analysed for 31 chemical and physical properties. Legume species present at each site were recorded and a presence/absence matrix was generated. Cluster analysis and discriminant function analysis (DFA) were run to group the sites based on overall similarity in edaphic characteristics and to identify the soil parameters contributing towards discriminating the groups. Canonical correspondence analysis (CCA) was used to test for a correlation between legume species compositions and edaphic factors. The strength of the association between legume species and site

groupings based on edaphic properties was assessed using indicator species analysis.

### **Important findings**

Based on similarity in overall soil characteristics, the sites formed three clusters: one comprising sites of sandstone geology, one with dune sand sites and the third cluster comprising sites of both shale and granite geologies (hereafter referred to as soil types). The DFA confirmed the distinctness of these clusters and the CCA showed a significant correlation between legume species composition and edaphic factors. The key edaphic parameters were clay content, iron (Fe), potassium (K), sulphur (S) and zinc (Zn). These findings reveal that the Cape Peninsula is edaphically heterogeneous and edaphically distinct habitats contain discrete legume species assemblages that can be distinguished by unique indicator species. Furthermore, multiple soil parameters, rather than a single parameter, are involved. Therefore, edaphic factors play a significant role in driving the distribution of legume species in the Cape Peninsula and discrete legume species assemblages occupy distinct habitats.

**Keywords:** Fabaceae, Core Cape Subregion, Cape Peninsula, edaphic factors, biogeography

Received: 15 March 2016, Revised: 17 March 2017, Accepted: 12 April 2017

## INTRODUCTION

Legumes have a global distribution (LPWG 2017; Schrire *et al.* 2005), but they mostly occur in seasonally dry tropical forests

and semi-arid temperate shrublands (Lavin and Matos 2008). Their success in these harsh environments has been linked with their high leaf nitrogen (N) content (McKey 1994), a trait likely evolved in ancestral legumes, which occupied

nutrient-rich habitats (Herendeen *et al.* 1992). Leaf nitrogen content is positively correlated with photosynthetic capacity (Chen *et al.* 2015; Tuohy *et al.* 1991), and thus facilitates rapid growth and opportunistic production of short-lived leaves, allowing legumes to capitalize on brief periods of resource availability (McKey 1994). Nitrogen fixation (N-fixation) is hypothesized to have evolved later as legumes spread into low-nitrogen habitats (McKey 1994). This hypothesis has not yet been tested, but studies indicate that legumes radiated rapidly, shortly after their origin, 60 million years ago (Bell *et al.* 2010), and N-fixation evolved multiple independent times during this radiation (Doyle 2011).

Considering their adaptation for harsh environments, it is not surprising that legumes are the second most species-rich lineage in the Core Cape Subregion (CCR, also known as the Cape Floristic Region) of South Africa (Manning and Goldblatt 2012). The CCR is one of the world's floristically rich regions (Cox 2001; Takhtajan 1986), with a Mediterranean type of climate (Goldblatt and Manning 2002). About 80% of its soils are derived from quartzitic sandstone and are mostly acidic and nutrient-poor, while some limestone, shale and granite-derived soils which are slightly higher in nutrients, especially nitrogen and phosphorus (P), constitute ~20% of the CCR (Campbell 1986; Cramer *et al.* 2014; Rebelo *et al.* 2006). Most CCR legumes are post-fire ephemerals (Le Maitre and Midgley 1992), belonging to the predominantly N-fixing subfamily Papilionoideae (Sprent 2007; Sprent *et al.* 2017). They also form cluster roots and symbiotic associations with mycorrhiza to facilitate P-uptake (Allsopp and Stock 1993; Lambers *et al.* 2010; Lamont 1982; Maistry *et al.* 2016).

Surprisingly, most vegetation units of the CCR are dominated by Proteaceae, Ericaceae and Restionaceae species, and not the species-rich Leguminosae. This might be due to the fact that 75% of CCR legumes are short-lived reseeders (Le Maitre and Midgley 1992) that emerge shortly after fire and gradually disappear with time after the fire (Kruger 1983). Nevertheless, the legume reseeders have patchy distributions, whereas the less prevalent resprouters have a more widespread distribution (Schutte *et al.* 1995). Furthermore, distinct legume species assemblages tend to occupy particular niches at any given locality. Drivers of these distribution patterns are not yet fully understood. Dispersal capability, climate, rhizobia specificity (see Lemaire *et al.* 2015) and edaphic factors are some likely drivers.

Reseeders tend to grow rapidly in order to complete their life cycle within a short space of time. Since this lifestyle comes at a high nutritional cost (Bell 2001; Lambers *et al.* 2008), reseeders tend to require higher nutrient levels than resprouters (Power *et al.* 2011). While N may not be limiting for the N-fixing legumes, other nutrients, like P, may still be limiting (Cramer 2010; Maistry *et al.* 2013; Power *et al.* 2010). Therefore, emerging shortly after fire allows legumes to capitalize on flushes of nutrients released after the fire event (Brown and Mitchell 1986; Certini 2005; Dean *et al.* 2015). Considering the patchy distribution of nutrient-enriched sites

following a fire event (Stock and Lewis 1986) and that species thrive on soils that meet their nutritional requirements (Richards *et al.* 1997a), the legumes might be tracking these nutrient-enriched patches, hence their patchy distributions. Could the tendency of having discrete legume assemblages occupying distinct habitats also be due to edaphic factors?

Pearson and Dawson (2003) conceptualized a hierarchical framework of the importance of various biotic and abiotic factors in driving the distribution of species at varying spatial scales. Climate has the strongest influence at the global, continental and regional (i.e. >200 km) scales; edaphic conditions are more influential at the local and site level (1–10 km and 10–1000 m, respectively), while biotic interactions are strongest from the local to the micro (<10 m) scales (Pearson and Dawson 2003). Several studies on plants and animals support this framework (Eiserhardt *et al.* 2011; Luoto *et al.* 2006; Pearson *et al.* 2004; Thuiller *et al.* 2004). However, deviations from this framework have also been reported. For example, Araújo and Luoto (2007) found that although climate was the major determining factor in the distribution of European Apollo butterflies, biotic interactions also had a significant role. Others include Heikkinen *et al.* (2007) and Austin and van Niel (2011). Therefore, considering the uniqueness of the CCR regarding its species richness and ecological heterogeneity, it is necessary to test if the framework applies to the region.

Soil is the foremost substrate that anchors most land plants and from which they obtain water and nutrients. Several studies attest to the significance of edaphic factors in the biogeography of plants (Bertrand *et al.* 2012; Clark *et al.* 1999; Dubuis *et al.* 2013; Hall *et al.* 2004; John *et al.* 2007; Richards *et al.* 1997b; Soares *et al.* 2015). However, the key edaphic parameters that drive species distributions tend to vary with the ecosystem and taxa involved. Soils of the CCR are generally oligotrophic, and P is considered to be a limiting nutrient (Cramer 2010; Lambers *et al.* 2010; Stock and Verboom 2012). Low P levels are due to the main parent geological material's (Table Mountain sandstone) low P content of ~300  $\mu\text{g g}^{-1}$  (Lambers *et al.* 2010; Marchant and Moore 1978). Furthermore, most of the P is often bound to iron (Fe), calcium (Ca) and aluminium (Al) cations (Mitchell *et al.* 1984; Payn and Clough 1989; Witkowski and Mitchell 1987) and thus not readily available for plants. Therefore, P and these cations might be important in driving species distributions in the CCR. Since most CCR legumes are capable of N-fixation (Lemaire *et al.* 2015), N may not be limiting legume distributions in the CCR.

Edaphic heterogeneity is often cited as one of the main drivers of species distributions in the CCR, but the evidence to support such is very sparse. The only CCR studies that we are aware of are those of Esler and Cowling (1993), focusing largely on *Pteronia* species in the semi-arid karoo biome and Richards *et al.* (1997b), who focused on three pairs of Proteaceae species in the Agulhas Plains. Both studies focused on the role of competition between closely related species and

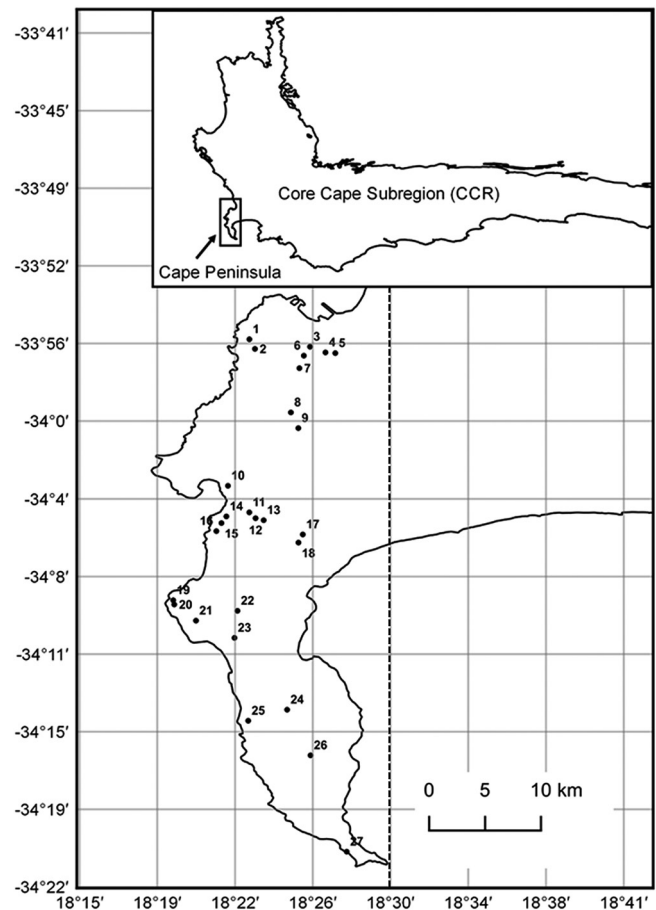
edaphic factors in structuring plant communities. The study by [Chimphango \*et al.\* \(2015\)](#), which showed that CCR legumes generally occupy nutrient-enriched islands within the landscape as compared to their surrounding non-legume vegetation, sheds light on the patchy distribution of legumes in the landscape. However, the question of having distinct legume assemblages being associated with distinct habitats remains unresolved. Hence, this study attempts to address it by identifying legume assemblages associated with distinct habitats and providing the link between legume distributions and edaphic factors, through sampling multiple legume communities across the edaphic gradients of the Cape Peninsula.

If edaphic factors drive the distribution of legumes in the Cape Peninsula, it should be possible to identify species or groups of species whose presence or absence in a given habitat signals the prevailing edaphic conditions of the habitat, *i.e.* indicator species ([Siddig \*et al.\* 2016](#)). Indicator species are species that can be used as ecological indicators of community types or environmental conditions due to their niche preferences ([Niemi and McDonald 2004](#)). Therefore, the aim of the study was to determine the role of edaphic factors in driving the distribution of legume species in the Cape Peninsula and to identify indicator species within the legume assemblages. Such knowledge will inform conservation planning aimed at curbing biodiversity loss and allows for prediction of how changes in nutrient deposition, nutrient cycling processes, fire regimes and climate, might impact the composition of the flora. Physical and chemical properties of soils associated with legume species assemblages were analysed to address the following key questions: (i) do edaphic factors (physical and chemical) influence the biogeography of Cape Peninsula legumes, if so, (ii) what are the key edaphic parameters driving the patterns, and (iii) can we identify indicator species for the various soil types in the Cape Peninsula? We hypothesized that the distributions of legume species are linearly related to edaphic factors and that the interaction of P, Al, Ca and Fe drives the legume species assemblages in the Cape Peninsula.

## MATERIALS AND METHODS

### Study site

The study was conducted in the Cape Peninsula, the area located within the southern core of the CCR, west of the 18°30'E longitude ([Fig. 1](#)), as defined by [Helme and Trinder-Smith \(2006\)](#). Topographically, it ranges from the low-lying Cape Flats to gentle slopes and hills, culminating in ridges and plateaus with a maximum altitude of 1113 m on Table Mountain ([Cowling \*et al.\* 1996](#)). The summits and upper slopes of the Cape Peninsula mountains are mostly associated with acidic and nutrient-poor sandstone-derived soils, whereas the lower slopes, which are associated with colluvium underlain by granite or shale, are less acidic and nutrient richer ([Cowling \*et al.\* 1996](#)). The coastal areas are mostly associated with alkaline and poorly consolidated sands ([Cowling \*et al.\* 1996](#)). The Cape Peninsula receives most of



**Figure 1:** location of the Cape Peninsula and the sampling sites within the CCR of South Africa. Grid squares of the Cape Peninsula are those located to the left of the 18°30' line (dashed line). Sampling sites are represented by the closed circles and their names are represented by the numbers as follows: 1 = Lions Head, 2 = Lions Head 1, 3 = Lower Devils Peak G, 4 = Lower Devils Peak UE, 5 = Lower Devils Peak RM, 6 = Tafelberg, 7 = Upper Devils Peak, 8 = Cecilia Ravine, 9 = Blue Gums, 10 = Blackburn Ravine, 11 = Disa Ridge 1, 12 = Disa Ridge, 13 = Silvermine Dam, 14 = Chapmans Peak foot, 15 = Lower Chapmans Peak, 16 = Chapmans Peak, 17 = Steenberg, 18 = Silvermine East, 19 = Kommetjie, 20 = Slangkop VS, 21 = Slangkop Egate, 22 = Redhill, 23 = Kleinplaas Dam, 24 = Smitswinkel Flats, 25 = Cape Point, 26 = Cape Point BF and 27 = Cape of Good Hope.

its rainfall during winter (June–August), while its summers are mostly dry. Annual precipitation ranges from 402 mm in Cape point to over 1000 mm in Maclear's Beacon, the highest point in the Cape Peninsula ([Cowling \*et al.\* 1996](#); [Higgins \*et al.\* 1999](#)). Annual precipitation varies considerably within very short distances owing to the highly heterogeneous altitude, aspect and topography ([Adamson and Salter 1950](#); [Cowling \*et al.\* 1996](#)).

Floristically, the Cape Peninsula has a total of 2285 species of vascular plants ([Trinder-Smith \*et al.\* 1996](#)), of which 158 species are endemic ([Helme and Trinder-Smith 2006](#)). The top five species-rich families in the Cape Peninsula are Asteraceae (286 species), Iridaceae (168 species), Fabaceae (162 species),

Poaceae (141 species) and Ericaceae (112 species) (Trinder-Smith *et al.* 1996). The major vegetation types that occur in the Cape Peninsula include fynbos, renosterveld, dune strandveld and afro-montane forest (Mucina and Rutherford 2006; Trinder-Smith 2006). Fynbos, the predominant vegetation type, is characterized by sclerophyllous shrubs and is dominated by Restionaceae, Proteaceae and Ericaceae species. Renosterveld occupies only 5% of the Cape Peninsula (Cowling *et al.* 1996) and is dominated by Asteraceae (especially Renosterbos: *Elytropappus rhinocerotis* (L.f) Less.), Poaceae species and various geophytic species. Dune strandveld consists of plants that grow on the inland edge of sandy beaches, stabilizing the soil. Forests are a minor component of the Cape Peninsula, restricted to moist valleys on eastern slopes and along river banks and consist of typical temperate tree genera such as *Podocarpus* L. (Campbell and Moll 1977).

### Sampling

Herbarium specimen records, dating as far back as the 1700s, were sourced from the Bolus Herbarium (BOL) and the Pretoria Computerised Information System (PRECIS) database of South African plants to determine the broad-scale distribution of legumes in the Cape Peninsula. These two sources carry the most comprehensive records of the CCR flora and the PRECIS database aggregates collections from all herbaria under the South African Biodiversity Institute (SANBI). However, the PRECIS records come georeferenced to 0.25° grid square, which is not ideal for this study, where finer scale geographic coordinates were required to accurately characterize species habitats. Therefore, all specimens obtained from the PRECIS database were further georeferenced using GEOLocate version 3.2.2 (Rios and Bart 2010) as the primary georeferencing tool and some detailed maps of the Cape Peninsula (Slingsby Maps). Most of the BOL specimens had already been accurately georeferenced, hence only those that had not been done were subjected to georeferencing.

To identify patterns of legume distributions within the Cape Peninsula (i.e. which areas have similar legume species composition) and areas of high legume species richness, the map of the Cape Peninsula was subdivided into 0.0625° grid squares. Thus, 19 such grid squares fall within the boundaries of the Cape Peninsula (Fig. 1). Using the georeferenced data, a matrix of presence (1) or absence (0) of all known (162) Cape Peninsula legume species in each of the 19 grid squares was generated. The matrix was analysed by cluster analysis (clustering algorithm: UPGMA, distance: Raup-Crick) using the program PAST version 3.12 (Hammer *et al.* 2001).

Results of the cluster analysis informed our field sampling, which sought to capture the diversity in legume species assemblages, geological substrates, altitude, aspect and topography of the Cape Peninsula. Accordingly, at least one site was selected from each cluster, and where members of the same cluster fell into different geological substrate types (inferred from overlaying a geology layer on the grids), at least one grid per geological type was sampled. Overall, a total of 27

legume sites, representing 12 of the 19 grid squares were sampled (Fig. 1). During fieldwork, each site was surveyed for legume species present through transect walks within 20 × 20 m quadrats. All legume species observed at each site were recorded and voucher specimens collected. Soil samples (three replicates) were randomly taken from the top layer (10–15 cm depth) of the soil profile by means of a soil corer or garden trowel. At the laboratory, the soil samples were air-dried at room temperature, all plant debris and roots removed and sieved through a 2 mm mesh.

### Analysis of chemical and physical properties of soils

The dried soil samples were analysed for 31 physical and chemical characteristics, including macro- and micro-elements, beneficial and toxic elements. These were: C, Ca, K, Mg, N (total N and NH<sub>4</sub>-N), P (total P and P Bray II) and S (i.e. macronutrients—Maathuis 2009); Cu, Fe, Mn, Mo, Ni, Zn (micronutrients—Hänsch and Mendel 2009); Al, Co, Ga, Na, Se, Si, V (beneficial elements—Pilon-Smits *et al.* 2009; Simon *et al.* 1989); As, Cd, Hg, Pb, Sn (toxic elements—Nagajyoti *et al.* 2010); as well as pH, sand, silt and clay contents. Each soil sample was partitioned into three portions. One portion was sent to the Elsenberg Laboratory of the Western Cape Government's Department of Agriculture for the analysis of ammonium (NH<sub>4</sub>-N), calcium (Ca), carbon (C), clay, magnesium (Mg), pH, potassium (K), sand, silt and sodium (Na) content. Water-soluble concentrations of the elements (i.e. available to plants) were measured for Ca, K, Mg and Na. Another portion was sent to Bemlab (Somerset West, South Africa) for the analysis of total P, P Bray II (available for plants) and total N. The final portion was crushed into a fine powder using mortar and pestle and analysed for total concentrations of the following 22 chemical elements: Al, As, Ca, Cd, Co, Cu, Fe, Ga, Hg, K, Mg, Mn, Mo, Ni, P, Pb, S, Se, Si, Sn, V and Zn using an X-ray fluorescence spectrophotometer.

### Determination of the species composition of sites

Although herbarium specimen records were initially used to identify the broad-scale legume species assemblages in the Cape Peninsula to locate sampling areas, such records are subject to spatial biases and georeferencing errors (Stropp *et al.* 2016), which compound an accurate determination of the fine-scale species composition of an area. Therefore, to compile a matrix of legume species compositions of the sampled sites, we used the data collected during our own field surveys. Each species was scored for presence (1) or absence (0) in each of the sites to develop a presence/absence matrix.

### Statistical analyses

The soil data were natural log-transformed where necessary to achieve normality. A hierarchical cluster analysis (clustering algorithm: Ward's, distance: Euclidean) of all soil parameters was conducted to group the sites based on overall similarity of soil characteristics. A canonical discriminant function

analysis (DFA) was used to determine the soil parameters that discriminate between the groups established from the cluster analysis. A stepwise DFA, whereby the variables are entered into the model one after the other, each time choosing the variable that maximizes the discriminatory power of the model, was used. Additionally, analysis of variance (ANOVA) was performed on the individual variables that had higher discriminatory contributions (based on their standardized coefficients) in the DFA. Tukey honest significance difference tests were used to identify significantly different means ( $P < 0.05$ ). Correlations between total and plant available concentrations of Ca, K, Mg and P were determined using Pearson correlation analysis. Except for the DFA, which was performed in Statistica (Statistica version 13, StatSoft, Inc., Tulsa, OK, USA), all analyses were performed in R (R version 3.3.2, R-core team 2016).

To test for correlation between species composition and edaphic factors, canonical correspondence analysis (CCA) was employed. The CCA is a multivariate technique that uses gradients in environmental attributes to depict species' habitat preferences on ordination graphs in which the axes are linear combinations of the environmental factors (Ter Braak 1986). This technique was chosen over other ordination methods because of its robustness to skewed species distributions, unusual sampling approaches and cases where not all variables driving species composition are known (Palmer 1993). Analyses were performed using the 'vegan' package in R. To select variables that best explain the species–environment relationship while accounting for correlation among variables (multicollinearity), we performed stepwise variable selection using the 'ordstep' function and examined the variance inflation factors (VIFs) of the variables (Ter Braak 1987). Variables with VIF > 10 were excluded from the analyses. Permutation tests (10 000 permutations) were run to evaluate the statistical significance of the CCA model, its terms (soil variables) and the CCA axes. Preliminary analyses showed that sites from dune sands had exceptional values for most variables, which skewed the rest of the sites in the CCA analyses, thus they were excluded from subsequent analyses. Additionally, three granite–shale sites were ordinated separately from the main group of granite–shale sites, but closer to the sandstone sites. Thus, additional cluster analysis, DFA and CCA were performed on the three granite–shale sites and all the sandstone sites to determine whether there were any distinct groupings within this subset of the data.

Partial Mantel tests were conducted to test for spatial autocorrelation in the species–environment relationship using the 'ecodist' package (Goslee and Urban 2007) in R. The partial Mantel test examines the correlation between two distance matrices while controlling for a third one. Consequently, it is widely used to identify spatial autocorrelation in ecological studies where the effects of environmental variables on species composition are determined while taking into account geographic distance (Giraldo *et al.* 2016; Martiny *et al.* 2011; Talbot *et al.* 2014). The statistical significance of the partial Mantel test results was assessed using 10 000 permutations.

We determined the strength of association between species and groups of sites using indicator values, which measure the predictive value of a species as an indicator of particular site groups (De Cáceres *et al.* 2010). Analyses were performed in R using the 'indicspecies' package. The 'IndVal.g' function of 'multipatt' which corrects for unequal group sizes was employed. Permutation tests (10 000) were performed to assess the statistical significance of the association between species and groups of sites. We used the species presence and absence matrix for the 27 sites and grouped the sites based on their distinctness in overall edaphic characteristics as depicted by the results of the DFA performed on the soil data.

## RESULTS

### Legumes species composition of sites

The total number of legume species encountered during the surveys was 67, representing 22 genera (out of 25) and 9 tribes (out of 10). The list of species observed in each of the sampled sites is found in online supplementary Table S1. Most species had restricted distributions, e.g. 38% of the species occurred on only one site; 26% on two sites; 14% on three sites and 22% of the species occurred on more than three sites (Fig. 2a). A few widespread species encountered include; *Bolusafra bituminosa* Kuntze and *Aspalathus ericifolia* Willd. ex. Walp., both occurring on seven sites, *Othobium virgatum* C.H.Stirt. was recorded on 11 sites and *Psoralea pinnata* L. occurred on 12 sites (see online supplementary Table S1). A majority of the species (66%) occurred on only one soil type, 21% on two soil types and only 13% of the species occurred on three different soil types (Fig. 2b). Although four soil types were studied (i.e. dune sands, granite, sandstone and shale), no single species occurred on all of them.

### Chemical and physical properties of soils

The cluster analysis of the overall soil characteristics grouped the 27 sites into three main groups (see online supplementary Fig. S1). One group comprised 16 sites, which are sites associated with Table Mountain sandstone parent rock material. Another group comprised two sites that are associated with coastal dune sands, while the third group (nine sites) comprised sites from both shale and granite substrates (see online supplementary Fig. S1). Of the 31 soil characteristics studied, 10 had significant contributions towards the discriminant function, thus the 10 were used for the DFA of the three groups obtained from the cluster analysis. The standardized coefficients of the 10 variables on the first and second roots of the DFA are shown in online supplementary Table S2. Chi-square tests showed that the two roots (discriminant functions) were significant ( $P < 0.001$ ). Root 1, which explained 69.09% of the variance in the model, was strongly influenced by Na, Fe, Al and K, in descending order, while for Root 2, which explained 30.91% of the variance, C and P were the most important elements (see online supplementary Table S2). A scatterplot of the canonical scores confirmed the

separation of the three groups along both roots (see online supplementary Fig. S2). In terms of univariate differences between the groups, sites from granite–shale substrates had

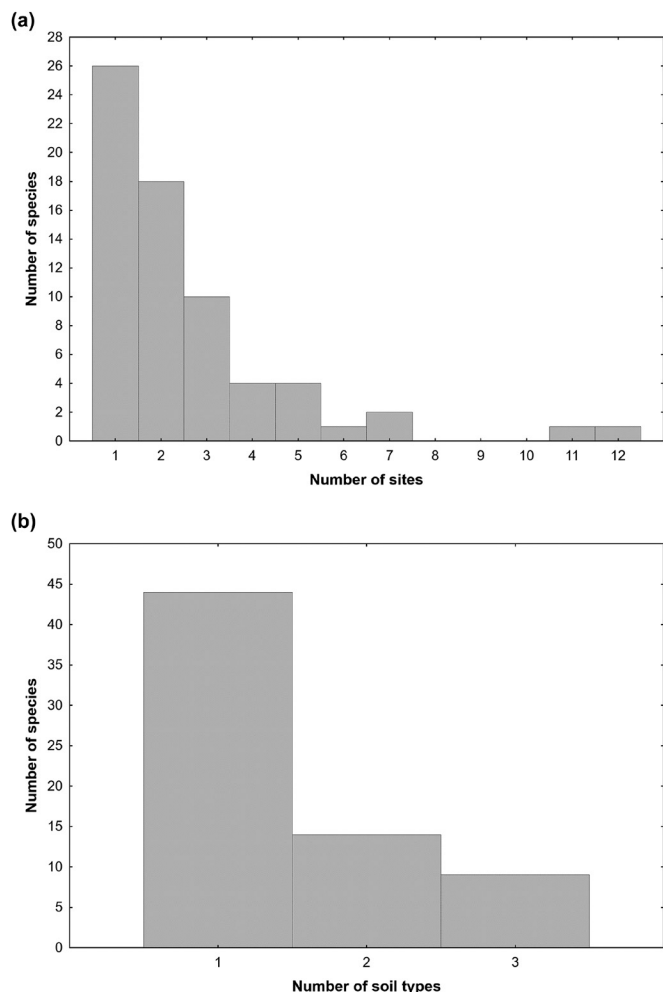
significantly higher Al, Fe, K, Mn and N, but lower sand content than the sites from sandstone and dune sands (Table 1). In contrast, the sites from dune sands had higher concentrations of P Bray II and Na than sites from the other two soil types (Table 1).

An assessment of correlation between total concentrations of Ca, K, Mg and P, as measured using the X-ray fluorescence technique and their corresponding plant available portions, showed a positive relationship: Pearson *r* values = 0.82, 0.85, 0.54 and 0.81, respectively ( $P < 0.0001$ ).

### Correlation between legume species composition and edaphic factors

Soil characteristics that best explained the species–environment relationship after removing collinear variables were clay content, Fe, K and Zn, hence they were used for the CCA. The CCA confirmed a significant linear relationship between species composition and soil characteristics (permutation  $F = 1.910$ ,  $P = 0.001$ ). The first and second canonical axes, which were both significant ( $P < 0.001$ ), jointly explained 82% of the constrained inertia (i.e. variance explained by the soil variables). The sites formed two main groups which separated along CCA1, one comprised six granite–shale sites and the other comprising three granite–shale sites ordinated next to the sandstone sites (Fig. 3a). The six granite–shale sites were associated with higher values of all four key soil characteristics. Legume species were distributed throughout all four quadrats of the first two CCA axes (Fig. 3b), indicating distinct edaphic optima for the different legume species assemblages.

Due to the clustering of three granite–shale sites with the sandstone sites (Fig. 3a), separate cluster analysis, DFA and CCA were performed on this group to determine whether there were any distinct sub-groups within it. The cluster analysis recovered two main clusters of the sandstone sites (denoted Sandstone\_1 and Sandstone\_2) and one cluster containing all three granite–shale sites and one sandstone site (denoted Granite–shale\_1) (see online supplementary Fig. S3). Six soil variables (C, Fe, K, N, P Bray II and pH) contributed significantly to the discrimination of the three clusters (see online

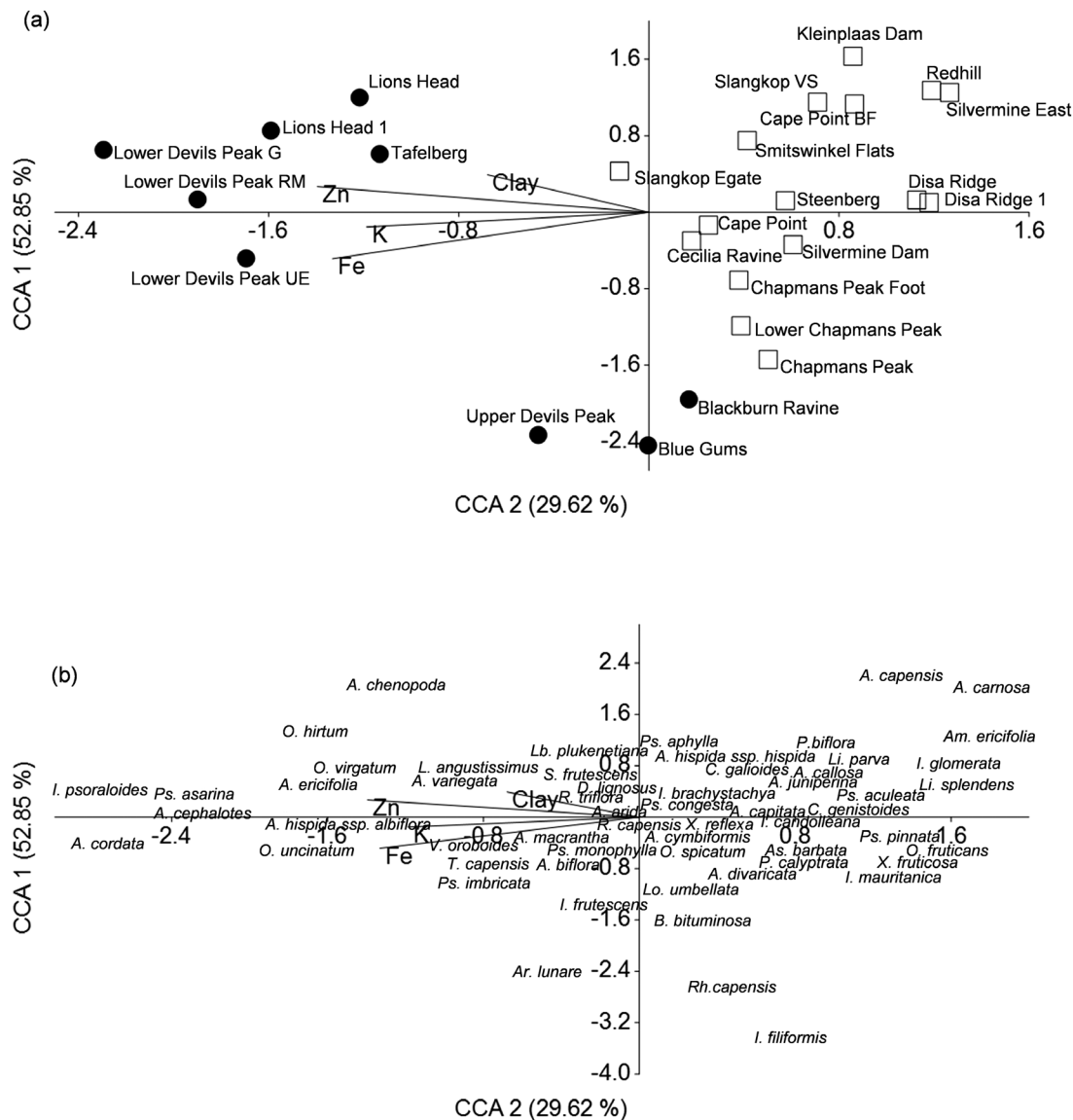


**Figure 2:** bar plot of the total numbers of legume species encountered in different numbers of sites (a) and soil types (b).

**Table 1:** mean  $\pm$  SE of the concentrations of 10 variables separating the three main soil groups associated with the 27 sites

Soil variable	Granite–shale	Sandstone	Dune sand
Al (%)	8.632 $\pm$ 0.468 <sup>a</sup>	2.174 $\pm$ 0.292 <sup>b</sup>	2.633 $\pm$ 0.876 <sup>b</sup>
P (%)	0.048 $\pm$ 0.007 <sup>a</sup>	0.027 $\pm$ 0.004 <sup>b</sup>	0.155 $\pm$ 0.013 <sup>c</sup>
Mn (%)	0.049 $\pm$ 0.004 <sup>a</sup>	0.003 $\pm$ 0.002 <sup>b</sup>	0.001 $\pm$ 0.007 <sup>b</sup>
Fe (%)	3.035 $\pm$ 0.154 <sup>a</sup>	0.342 $\pm$ 0.096 <sup>b</sup>	0.109 $\pm$ 0.287 <sup>b</sup>
P Bray II (mg kg <sup>-1</sup> )	8.591 $\pm$ 3.166 <sup>a</sup>	2.083 $\pm$ 1.974 <sup>a</sup>	77.878 $\pm$ 5.923 <sup>b</sup>
Na (mg kg <sup>-1</sup> )	79.381 $\pm$ 14.428 <sup>a</sup>	33.556 $\pm$ 8.998 <sup>b</sup>	222.167 $\pm$ 26.993 <sup>c</sup>
K (mg kg <sup>-1</sup> )	160.714 $\pm$ 8.337 <sup>a</sup>	35.037 $\pm$ 5.199 <sup>b</sup>	23.333 $\pm$ 15.597 <sup>b</sup>
N (%)	0.123 $\pm$ 0.01 <sup>a</sup>	0.078 $\pm$ 0.006 <sup>b</sup>	0.033 $\pm$ 0.018 <sup>b</sup>
C (%)	2.218 $\pm$ 0.293 <sup>a</sup>	2.083 $\pm$ 0.183 <sup>a</sup>	1.467 $\pm$ 0.548 <sup>a</sup>
Sand (%)	86.952 $\pm$ 0.715 <sup>a</sup>	94.704 $\pm$ 0.446 <sup>b</sup>	96.0 $\pm$ 1.338 <sup>b</sup>

Values with different letters are significantly different at  $P < 0.05$ .



**Figure 3:** (a) species-conditional biplot based on a CCA of soil characteristics and legume species composition (for all sites except the dune sand sites), showing the ordination of legume sites along the first two CCA axes. Sandstone sites are represented by open squares, while the closed circles represent granite–shale sites. Soil characteristics are indicated by lines whose lengths indicate the strength of their correlations with the ordination axes. (b) species-conditional biplot based on a CCA of soil characteristics and legume species composition (for all sites except the dune sand sites), showing the ordination of legume species along the first two CCA axes. Soil characteristics are indicated by lines, whose lengths indicate the strength of their correlations with the ordination axes. Genera are abbreviated as follows: *A* = *Aspalathus*, *Am* = *Amphithalea*, *Ar* = *Argyrolobium*, *B* = *Bolusafra*, *C* = *Cyclopia*, *D* = *Dipogon*, *I* = *Indigofera*, *L* = *Lotus*, *Le* = *Lessertia*, *Li* = *Liparia*, *Lo* = *Lotononis*, *O* = *Otholobium*, *P* = *Podalyria*, *Ps* = *Psoralea*, *R* = *Rafnia*, *Rh* = *Rhynchosia*, *T* = *Tephrosia*, *V* = *Virgilia*, *W* = *Wiborgia*, *X* = *Xiphotheca*.

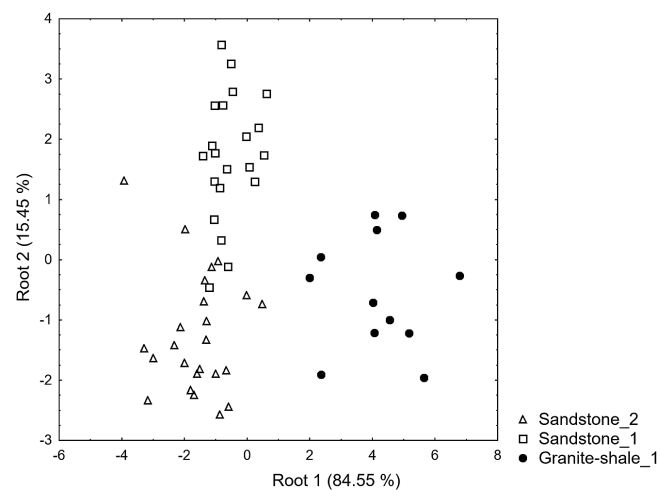
supplementary Table S3) and the DFA confirmed significant differences between the clusters ( $P < 0.0001$ ). Their ordination along the two roots is presented in Fig. 4. Root 1 (accounting for 84.55% of the variance) was strongly associated with Fe and Root 2 (15.45%) was strongly associated with C (see online supplementary Table S3). For the univariate ANOVA among the groups, Fe and P Bray II were significantly higher in the granite–shale sites than the two groups of sandstone sites (Table 2). The two groups of sandstone sites differed significantly in their C, K and N concentrations (Table 2).

Iron, K and S were the variables that contributed significantly to the species–environment relationship for this subset of the data. The CCA showed a significant correlation between species composition and soil characteristics (permutation  $F = 1.764$ ,  $P = 0.001$ ). The first two axes explained 69.29% ( $P = 0.001$ ) and 30.69% ( $P = 0.002$ ) of the constrained inertia, respectively. The Granite–shale\_1 sites clustered together and separately from the sandstone sites, and were associated with higher Fe and K (see online supplementary Fig. S4a). With the exception of two sites, the sites of

the Sandstone\_2 group formed a distinct cluster, while those from the Sandstone\_1 group had no distinct grouping (see online supplementary Fig. S4a). The species were distributed on all four quadrants of the CCA plot with various assemblages having distinct edaphic optima (see online supplementary Fig. S4b).

The partial Mantel test on species composition vs. geographic distance while accounting for edaphic factors was not significant (Mantel  $r = 0.090$ ,  $P = 0.097$ ), whereas the correlation between edaphic factors and species composition while taking into account geographic distance was significant (Mantel  $r = 0.280$ ,  $P = 0.0001$ ), indicating that there was no significant spatial autocorrelation in the data. Similar results were obtained for the analysis of the subset of the data (comprising three granite–shale and all sandstone sites). Thus, the observed correlations in the species–environment relationships were not compounded by spatial autocorrelation.

Based on the results of the DFA, the following groups were used for indicator species analysis: granite–shale, dune sand (see online supplementary Fig. S2), Sandstone\_1, Sandstone\_2 and Granite–shale\_1 (Fig. 4). A total of 13 species were identified as significant indicator species, three for Sandstone\_2,



**Figure 4:** scatterplot of the canonical scores for the two discriminant functions of the subset of the data comprising three granite–shale sites and all the sandstone sites.

two species for dune sands, two for Granite–shale\_1 and six species for the main granite–shale sites (Table 3). No species were associated with more than one group and only one site group (Sandstone\_1) did not have any significant indicator species assigned to it.

## DISCUSSION

The results of the DFA on the soil data showed that soils of the Cape Peninsula differ in physical and chemical properties that are known to influence plant growth and metabolism. Although not all soil characteristics were analysed for water-soluble portions (i.e. plant available concentrations), the findings of significant positive correlations between total and available concentrations of four elements (where concentrations were measured from both total and plant available forms) validate inferences of the role of the studied edaphic factors in driving legume species distributions. The results indicate that the soils differ in their physical and chemical properties according to their parent geology, except for the granite and shale, which were grouping together. Additionally, variation exists within the different soil types as indicated by the observation of a further significant split within the sandstone and granite–shale sites (Fig. 4). Importantly, this variation within soil types can be associated with varying species composition on the landscape as was shown in the present study.

The CCA confirmed a linear relationship between legume species composition of sites and edaphic factors, and this relationship was not due to spatial autocorrelation, as indicated by the results of the partial Mantel tests. Thus, habitats that differ in their soil characteristics are associated with different legume species assemblages. This is consistent with broad vegetation patterns in the CCR, whereby landscapes of a particular geology are occupied by specific taxa. For example, Ericaceae, Proteaceae and Restionaceae tend to be restricted to sandstone-derived soils, while *Mesembryanthemum*, *Oxalis* and Scrophulariaceae mostly occur on shale (Linder 2003). This suggests that edaphic factors play a significant role in determining what species can occupy particular habitats in the CCR. Therefore, the findings confirm our hypothesis that edaphic factors influence the biogeography of legume species in the Cape Peninsula and it corroborates the hierarchical

**Table 2:** mean  $\pm$  SE of the concentrations of the six key variables separating the three main soil groups associated with the sandstone and three granite–shale sites only

Soil variable	Sandstone_1	Sandstone_2	Granite–shale
Fe (%)	0.275 $\pm$ 0.118 <sup>a</sup>	0.129 $\pm$ 0.110 <sup>a</sup>	1.508 $\pm$ 0.156 <sup>b</sup>
K (mg kg <sup>-1</sup> )	46.286 $\pm$ 4.099 <sup>a</sup>	22.583 $\pm$ 3.834 <sup>b</sup>	52.167 $\pm$ 5.423 <sup>a</sup>
N (%)	0.085 $\pm$ 0.008 <sup>a</sup>	0.055 $\pm$ 0.008 <sup>b</sup>	0.117 $\pm$ 0.011 <sup>a</sup>
C (%)	2.632 $\pm$ 0.263 <sup>a</sup>	1.120 $\pm$ 0.246 <sup>b</sup>	2.956 $\pm$ 0.347 <sup>a</sup>
P Bray II (mg kg <sup>-1</sup> )	1.410 $\pm$ 0.372 <sup>a</sup>	1.597 $\pm$ 0.348 <sup>a</sup>	4.522 $\pm$ 0.492 <sup>b</sup>
pH (KCl)	3.790 $\pm$ 0.166 <sup>a</sup>	4.196 $\pm$ 0.155 <sup>a</sup>	4.142 $\pm$ 0.220 <sup>a</sup>

Values with different letters are significantly different at  $P < 0.05$ .



**Table 3:** results of indicator species analysis for the legume sites of the Cape Peninsula

Species	Group	A	B	IndVal	P-value
<i>Amphithalea ericifolia</i> (L.) Eckl. & Zeyh.	Sandstone_2	0.889	1.00	0.943	0.00001
<i>Aspalathus capensis</i> (Walp.) R. Dahlgren	Sandstone_2	1.00	0.875	0.935	0.00001
<i>Aspalathus carnosa</i> Eckl. & Zeyh.	Sandstone_2	1.00	1.00	1.00	0.0002
<i>Otholobium bracteolatum</i> (Eckl. & Zeyh.) C. H. Stirt.	Dune sand	1.00	1.00	1.00	0.0392
<i>Psoralea repens</i> P. J. Bergius	Dune sand	1.00	1.00	1.00	0.0392
<i>Argyrolobium lunare</i> Druce	Granite–shale_1	0.857	1.00	0.926	0.0012
<i>Indigofera filiformis</i> Thunb.	Granite–shale_1	0.800	1.00	0.894	0.0021
<i>Aspalathus cordata</i> (L.) R. Dahlgren	Granite–shale	1.00	1.00	1.00	0.0001
<i>Indigofera psoraloides</i> L.	Granite–shale	1.00	1.00	1.00	0.0001
<i>Otholobium hirtum</i> (L.) C. H. Stirt.	Granite–shale	1.00	1.00	1.00	0.0001
<i>Psoralea asarina</i> (P. J. Bergius) T. M. Salter	Granite–shale	1.00	1.00	1.00	0.0001
<i>Aspalathus hispida</i> ssp. albiflora (Eckl. & Zeyh.) R. Dahlgren	Granite–shale	1.00	0.833	0.913	0.0004
<i>Aspalathus chenopoda</i> L.	Granite–shale	0.800	1.00	0.894	0.0015

For each species, we present the probability that the surveyed site belongs to the target site group (A), the probability of finding the species in sites belonging to the site group (B), the Indicator value (IndVal) which is a product of A and B (De Cáceres *et al.* 2010) and the statistical significance of the association (*P*-value).

framework of Pearson and Dawson (2003), which highlights edaphic factors as important for species distributions at such local spatial scales. Similar findings have been reported from several ecosystems, e.g. temperate forests (Bertrand *et al.* 2012), tropical rain forests (Clark *et al.* 1999; John *et al.* 2007), Alps (Dubuis *et al.* 2013), Mediterranean shrublands (Richards *et al.* 1997b) and savannas (Soares *et al.* 2015). Therefore, studies on the distributions of species at local spatial scales need to consider the effects of soils.

The observation of several species being restricted to fewer sites or soil types (Fig. 2) and the observation of distinct optima for the various legume species (Fig. 3b; see online supplementary Fig. S4b) indicate that the species differ in their edaphic requirements. These differences in edaphic preferences might allow the legumes to take advantage of the heterogeneous nature of the CCR's edaphic environment through diversifying their niches, thereby promoting speciation. This could account for the high legume species richness of the CCR and it would be consistent with the hypothesis that the high ecological heterogeneity of the Cape Peninsula promotes higher beta diversity, thereby leading to its high species richness (Simmons and Cowling 1996).

The main soil characteristics involved in driving the distribution of the Cape Peninsula legumes were clay, Fe, K, S and Zn (Fig. 3; see online supplementary Fig. S4). The involvement of multiple elements underscores the idea that focusing on a single soil parameter for species distributions may be misleading (Chimphango *et al.* 2015). These soil parameters are known for their roles in plant growth and their involvement in driving plant species distributions is well documented. Chimphango *et al.* (2015) reported K among the nutrients which were higher in legume patches than nearby non-legume vegetation. In this study, K was over four times higher in

the granite–shale sites than the sandstone or dune sand sites (Table 1). Moreover, the CCA separated granite–shale sites from sandstone sites (Fig. 3; see online supplementary Fig. S4) indicating that species on these habitats have different nutritional optima. Considering that four of the six indicator species for the granite–shale group (Table 3) occur exclusively on granite–shale habitats (see online supplementary Table S1), it is possible that such species cannot survive on the low K sandstone and dune sand habitats, hence they are restricted to the high K shale and granite habitats. Therefore, K must be important for legume distribution in the CCR.

Sulphur is a macronutrient involved in the synthesis of proteins, vitamins, chlorophyll and defence compounds against biotic and abiotic stress (Rausch and Wachter 2005). Likewise, Zn is an essential component of many proteins in plants and while its deficiency may inhibit plant growth and metabolism, excess amounts of Zn can be toxic to plants (Broadley *et al.* 2007). Therefore, the involvement of S and Zn in the legume distributions suggests that variations in their availability in the landscape could impact species distributions. Clay content is known to influence a number of soil properties, e.g. organic matter content, water retention and infiltration capacity, thus its involvement here highlights its potential significance for species distributions in the CCR.

We hypothesized that the interaction of P with Al, Ca and Fe are the main drivers of legume distributions in the Cape Peninsula. The Fe content of granite–shale sites was up to 10 times higher than that of sandstone and dune sand sites (Tables 1 and 2) and it was the strongest driver of the legume species composition–soil relationship in the CCA (Fig. 3; see online supplementary Fig. S4). Its involvement might be linked to its tendency to bind P (the main limiting nutrient), making it unavailable to plants (Witkowski and Mitchell 1987). Apart

from its role in respiration, photosynthesis, hormone structure and DNA synthesis (Graziano and Lamatina 2005; Jeong and Guerinot 2009; Rout and Sahoo 2015), Fe is needed for nodulation and N-fixation (Rotaru and Sinclair 2009; Tang et al. 1990). Considering that CCR soils are poor in N (Kruger 1983; Stock and Lewis 1986; Witkowski and Mitchell 1987), whereas legumes have a higher N requirement (McKey 1994), N-fixation must be the primary source of N for the legumes. Indeed, a majority of the observed native legumes in the CCR are N-fixing (Lemaire et al. 2015). Therefore, because of its involvement in N-fixation, Fe must be important for legumes in the CCR, thus influencing their distribution. However, in excess amounts, Fe can be toxic to plants through generating oxidative stress (Kampfenkel et al. 1995), thus it could also be acting as an environmental filter, allowing only those species that can tolerate or counteract its toxicity to survive in a given habitat.

The indicator species analysis identified 13 significant associations of legume species with particular groups of sites based on their soils' chemical and physical properties. This accentuates the significance of the correlation between edaphic factors and legume species composition of sites. Thus, by observing these indicator species on a given site, one can potentially predict the edaphic conditions of the site and similarly, given the list of species occurring at a site, one can predict the prevailing edaphic conditions of the site. This observation holds true outside the Cape Peninsula where a number of the indicator species occupy sites with (potentially) similar edaphic conditions. For example, *O. bracteolatum* is widespread within the CCR where it is found in the strandveld, a vegetation type characterizing the dune sand ecosystem. Thus, our results on indicator species have broader utility. Indicator species are important for conservation planning in that their presence or absence in a particular habitat could guide delineation of ecoregions or provide a signal of changes in the state of the environment such as nutrient deposition from air pollution. Therefore, our findings provide a basis for further studies incorporating more environmental attributes to strengthen our understanding of the interplay between the environment and community assembly processes in the CCR.

## CONCLUSION

The study has shown that the Cape Peninsula is edaphically heterogeneous and that differences in soil characteristics of sites are associated with differences in legume species assemblages. In addition, multiple soil parameters rather than a single soil parameter are involved in driving the legume species distributions. The study also showed that some legume species can serve as indicator species for the edaphic conditions of the sites they occupy. Therefore, soil chemical and physical factors contribute significantly towards the biogeography of legumes in the Cape Peninsula. Considering that the Cape Peninsula is a microcosm of the CCR in terms of legume species diversity and edaphic habitat types, we predict that

similar results can be obtained with a further sampling of legumes across the CCR. These findings imply that conservation planning and studies seeking to predict future distributions of the legumes, such as those relating to the impacts of climate change, need to consider patterns and amounts of nutrient deposition that could affect the survival of some species in an area or their migration out of these areas. We recommend that further studies investigate the effects of climate, dispersal and biotic interactions in order to assess their relative contributions towards the biogeography of legumes in the CCR.

## SUPPLEMENTARY DATA

Supplementary data are available at *Journal of Plant Ecology* online.

## FUNDING

This work was supported by the National Research Foundation of South Africa (Grant number 81818: Biology of Cape Legumes to AMM).

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the Table Mountain National Parks for granting permission to conduct the research in protected areas under their management. Thanks to Mr Edward Chirwa, Ms Elelwani Nenzhelele and Mr Odwa Obose for assistance with fieldwork and some laboratory analyses. We would also like to thank three anonymous referees who made some vital comments on an earlier version of this manuscript.

*Conflict of interest statement:* None declared.

## REFERENCES

- Adamson RS, Salter TM (1950) *Flora of the Cape Peninsula*. Cape Town & Johannesburg: Juta & Co.
- Allsopp N, Stock WD (1993) Mycorrhizal status of plants growing in the Cape Floristic Region, South Africa. *Bothalia* **23**:91–104.
- Araújo MB, Luoto M (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecol Biogeogr* **16**:743–53.
- Austin MP, Van Niel KP (2011) Improving species distribution models for climate change studies: variable selection and scale. *J Biogeogr* **38**:1–8.
- Bell CD, Soltis DE, Soltis PS (2010) The age and diversification of the angiosperms re-revisited. *Am J Bot* **97**:1296–303.
- Bell DT (2001) Ecological response syndromes in the flora of southwestern Western Australia: fire resprouters versus reseeder. *Bot Rev* **67**:417–33.
- Bertrand R, Perez V, Gegout J (2012) Disregarding the edaphic dimension in species distribution models leads to the omission of crucial spatial information under climate change: the case of *Quercus pubescens* in France. *Glob Change Biol* **18**:2648–60.
- Broadley MR, White PJ, Hammond JP, et al. (2007) Zinc in plants. *New Phytol* **173**:677–702.

- Brown G, Mitchell D (1986) Influence of fire on the soil phosphorus status in sand plain lowland fynbos, south-western Cape. *S Afr J Bot* **52**:67–72.
- Campbell B (1986) Vegetation classification in a floristically complex area: the Cape Floristic Region. *S Afr J Bot* **52**:129–40.
- Campbell B, Moll E (1977) The forest communities of Table Mountain, South Africa. *Vegetatio* **34**:105–15.
- Certini G (2005) Effects of fire on properties of forest soils: a review. *Oecologia* **143**:1–10.
- Chen Z, Zha T, Jia X, *et al.* (2015) Leaf nitrogen is closely coupled to phenophases in a desert shrub ecosystem in China. *J Arid Environ* **122**:124–31.
- Chimphango SB, Potgieter G, Cramer MD (2015) Differentiation of the biogeochemical niches of legumes and non-legumes in the Cape Floristic Region of South Africa. *Plant Ecol* **216**:1583–95.
- Clark DB, Palmer MW, Clark DA (1999) Edaphic factors and the landscape-scale distributions of tropical forest trees. *Ecology* **80**:2662–75.
- Cowling RM, Rundel PW, Lamont BB, *et al.* (1996) Plant diversity in mediterranean-climate regions. *Trends Ecol Evol* **11**:362–6.
- Cox CB (2001) The biogeographic regions reconsidered. *J Biogeogr* **28**:511–23.
- Cramer MD (2010) Phosphate as a limiting resource: introduction. *Plant Soil* **334**:1–10.
- Cramer M, West A, Power S, *et al.* (2014) Plant ecophysiological diversity. In Allsopp N, Colville J, Verboom A (eds). *Fynbos: Ecology, Evolution, and Conservation of a Megadiverse Region*. England: Oxford University Press, 248–72.
- Dean S, Farrer EC, Menges ES (2015) Fire effects on soil biogeochemistry in Florida scrubby flatwoods. *Am Midl Nat* **174**:49–64.
- De Cáceres M, Legendre P, Moretti M (2010) Improving indicator species analysis by combining groups of sites. *Oikos* **119**:1674–84.
- Doyle JJ (2011) Phylogenetic perspectives on the origins of nodulation. *Mol Plant Microbe Interact* **24**:1289–95.
- Dubuis A, Giovanettina S, Pellissier L, *et al.* (2013) Improving the prediction of plant species distribution and community composition by adding edaphic to topo-climatic variables. *J Veg Sci* **24**:593–606.
- Eisnerhard WL, Svenning JC, Kissling WD, *et al.* (2011) Geographical ecology of the palms (Arecaceae): determinants of diversity and distributions across spatial scales. *Ann Bot* **108**:1391–416.
- Esler KJ, Cowling RM (1993) Edaphic factors and competition as determinants of pattern in South African karoo vegetation. *S Afr J Bot* **59**:287–95.
- Giraldo R, Caballero W, Camacho-Tamayo J (2016) Mantel test for spatial functional data. *AStA Adv Stat Anal*, 10.1007/s10182-016-0280-1.
- Goldblatt P, Manning JC (2002) Plant diversity of the Cape region of South Africa. *Ann Mo Bot Gard* **89**:281–302.
- Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *J Stat Softw* **22**, 10.18637/jss.v022.i07.
- Graziano M, Lamattina L (2005) Nitric oxide and iron in plants: an emerging and converging story. *Trends Plant Sci* **10**:4–8.
- Hall JS, Mckenna JJ, Ashton PMS, *et al.* (2004) Habitat characterizations underestimate the role of edaphic factors controlling the distribution of *Entandrophragma*. *Ecology* **85**:2171–83.
- Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological statistics software package for education and data analysis. *Palaentol Electron* **4**:9.
- Hänsch R, Mendel RR (2009) Physiological functions of mineral micronutrients (Cu, Zn, Mn, Fe, Ni, Mo, B, Cl). *Curr Opin Plant Biol* **12**:259–66.
- Heikkinen RK, Luoto M, Virkkala R, *et al.* (2007) Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecol Biogeogr* **16**:754–63.
- Helme N, Trinder-Smith T (2006) The endemic flora of the Cape Peninsula, South Africa. *S Afr J Bot* **72**:205–10.
- Herendeen PS, Crepet WL, Dilcher DL (1992) The fossil history of the Leguminosae: phylogenetic and biogeographical implications. In Herendeen PS, Dilcher DL (eds). *Advances in Legume Systematics 4*. Kew: Royal Botanic Gardens, 303–16.
- Higgins SI, Richardson DM, Cowling RM, *et al.* (1999) Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. *Conserv Biol* **13**:303–13.
- Jeong J, Guerinot ML (2009) Homing in on iron homeostasis in plants. *Trends Plant Sci* **14**:280–5.
- John R, Dalling JW, Harms KE, *et al.* (2007) Soil nutrients influence spatial distributions of tropical tree species. *Proc Natl Acad Sci USA* **104**:864–9.
- Kampfenkel K, Van Montagu M, Inze D (1995) Effects of iron excess on *Nicotiana plumbaginifolia* plants (implications to oxidative stress). *Plant Physiol* **107**:725–35.
- Kruger FJ (1983) Plant community diversity and dynamics in relation to fire. In Kruger FJ, Mitchell DT, Jarvus JUM (eds). *Mediterranean-type Ecosystems*. Berlin: Springer, 446–72.
- Labbers H, Brundrett MC, Raven JA, *et al.* (2010) Plant mineral nutrition in ancient landscapes: high plant species diversity on infertile soils is linked to functional diversity for nutritional strategies. *Plant Soil* **334**:11–31.
- Labbers H, Raven JA, Shaver GR, *et al.* (2008) Plant nutrient-acquisition strategies change with soil age. *Trends Ecol Evol* **23**:95–103.
- Lamont B (1982) Mechanisms for enhancing nutrient uptake in plants, with particular reference to Mediterranean South Africa and Western Australia. *Bot Rev* **48**:597–689.
- Lavin M, Matos AB (2008) The impact of ecology and biogeography on legume diversity, endemism, and phylogeny in the Caribbean region: a new direction in historical biogeography. *Bot Rev* **74**:178–96.
- Le Maitre D, Midgley J (1992) Plant reproductive ecology. In Cowling RM (ed). *The Ecology of Fynbos: Nutrients, Fire and Diversity*. Cape Town: Oxford University Press, 135–75.
- Lemaire B, Dlodlo O, Chimphango S, *et al.* (2015) Symbiotic diversity, specificity and distribution of rhizobia in native legumes of the Core Cape Subregion (South Africa). *FEMS Microbiol Ecol* **91**:1–17.
- Linder HP (2003) The radiation of the Cape flora, southern Africa. *Biol Rev Camb Philos Soc* **78**:597–638.
- LPWG (2017) A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. *Taxon* **66**:44–77.

- Luoto M, Heikkinen R, Pöyry J, et al. (2006) Determinants of the biogeographical distribution of butterflies in boreal regions. *J Biogeogr* **33**:1764–78.
- Maathuis FJ (2009) Physiological functions of mineral macronutrients. *Curr Opin Plant Biol* **12**:250–8.
- Maistry PM, Cramer MD, Chimphango SBM (2013) N and P colimitation of N<sub>2</sub>-fixing and N-supplied fynbos legumes from the Cape Floristic Region. *Plant Soil* **373**:217–28.
- Maistry PM, Muasya AM, Valentine AJ, et al. (2016) Closely related allopatric *Podalyria* species from the Core Cape Subregion differ in their mechanisms for acquisition of phosphorus, growth and ecological niche. *J Plant Ecol* **9**:451–563.
- Manning J, Goldblatt P (2012) *Plants of the Greater Cape Floristic Region 1: The Core Cape flora*. *Strelitzia* 29. Pretoria: South African National Biodiversity Institute.
- Marchant J, Moore A (1978) Geochemistry of the Table Mountain Group: II. Analysis of two suites of Western Graafwater Rocks. *Trans Geol Soc S Afr* **81**:353–59.
- Martiny JB, Eisen JA, Penn K, et al. (2011) Drivers of bacterial beta-diversity depend on spatial scale. *Proc Natl Acad Sci USA* **108**:7850–4.
- McKey D (1994) Legumes and nitrogen: the evolutionary ecology of a nitrogen-demanding lifestyle. In Sprent JI, McKey D (eds). *Advances in Legume Systematics 5: The Nitrogen Factor*. Kew: Royal Botanic Gardens, 211–28.
- Mitchell D, Brown G, Jongens-Roberts S (1984) Variation of forms of phosphorus in the sandy soils of coastal fynbos, south-western Cape. *J Ecol* **72**:575–84.
- Mucina L, Rutherford MC (2006) *The Vegetation of South Africa, Lesotho and Swaziland*. *Strelitzia* 19. Pretoria: South African National Biodiversity Institute.
- Niemi GJ, McDonald ME (2004) Application of ecological indicators. *Annu Rev Ecol Evol Syst* **35**:89–111.
- Palmer MW (1993) Putting things in even better order: the advantages of canonical correspondence analysis. *Ecology* **74**:2215–30.
- Payn T, Clough M (1989) Factors affecting soil phosphate distribution and reactions in Cape forest soils. *South African Forest J* **150**:24–31.
- Pearson RG, Dawson TP, Liu C (2004) Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography* **27**:285–298.
- Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecol Biogeogr* **12**:361–71.
- Pilon-Smits EA, Quinn CF, Tapken W, et al. (2009) Physiological functions of beneficial elements. *Curr Opin Plant Biol* **12**:267–74.
- Power SC, Cramer MD, Verboom GA, et al. (2010) Does phosphate acquisition constrain legume persistence in the fynbos of the Cape Floristic Region? *Plant Soil* **334**:33–46.
- Power SC, Cramer MD, Verboom GA, et al. (2011) Legume seeders of the Cape Floristic Region inhabit more fertile soils than congeneric resprouters—sometimes. *Plant Ecol* **212**:1979–89.
- R Core Team (2016) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> (13 December 2016, date last accessed).
- Rausch T, Wachter A (2005) Sulfur metabolism: a versatile platform for launching defence operations. *Trends Plant Sci* **10**:503–9.
- Rebello AG, Boucher C, Helme N, et al. (2006) Fynbos Biome. In Mucina L, Rutherford MC (eds). *The Vegetation of South Africa, Lesotho and Swaziland*. *Strelitzia* 19. Pretoria: South African National Biodiversity Institute, 55–206.
- Richards MB, Cowling RM, Stock WD (1997a) Soil factors and competition as determinants of the distribution of six fynbos Proteaceae species. *Oikos* **79**:394–406.
- Richards MB, Stock WD, Cowling RM (1997b) Soil nutrient dynamics and community boundaries in the fynbos vegetation of South Africa. *Plant Ecol* **130**:143–55.
- Rios NE, Bart HL (2010) *GEOLocate Version 3.2.2*. Belle Chasse, LA: Tulane University Museum of Natural History.
- Rotaru V, Sinclair TR (2009) Interactive influence of phosphorus and iron on nitrogen fixation by soybean. *Environ Exp Bot* **66**:94–9.
- Rout GR, Sahoo S (2015) Role of iron in plant growth and metabolism. *Rev Agricult Sci* **3**:1–24.
- Schrire B, Lewis G, Lavin M (2005) Biogeography of the Leguminosae. In Lewis G, Schrire BD, Lock M (eds). *Legumes of the World*. Kew: Royal Botanic Gardens, 21–54.
- Schutte A, Vlok J, Van Wyk B (1995) Fire-survival strategy – a character of taxonomic, ecological and evolutionary importance in fynbos legumes. *Plant Syst Evol* **195**:243–59.
- Siddig AAH, Ellison AM, Ochs A, et al. (2016) How do ecologists select and use indicator species to monitor ecological change? Insights from 14 years of publication in ecological indicators. *Ecol Indic* **60**: 223–30.
- Simmons M, Cowling R (1996) Why is the Cape Peninsula so rich in plant species? An analysis of the independent diversity components. *Biodivers Conserv* **5**:551–73.
- Simon L, Kiss F, Bako A, et al. (1989) Effect of gallium on photosynthetic pigments and peroxidase activity of *Chlorella pyrenoidosa*. *J Plant Nutr* **12**:1123–40.
- Soares MP, Reys P, Pifano DS, et al. (2015) Relationship between edaphic factors and vegetation in savannas of the Brazilian Midwest Region. *Rev Brasileira Ciência do Solo* **39**:821–29.
- Sprent JI (2007) Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. *New Phytol* **174**:11–25.
- Sprent JI, Ardley J, James EK (2017) Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytol* **215**:40–56.
- Stock W, Lewis O (1986) Soil nitrogen and the role of fire as a mineralizing agent in a South African coastal fynbos ecosystem. *J Ecol* **74**:317–28.
- Stock WD, Verboom GA (2012) Phylogenetic ecology of foliar N and P concentrations and N:P ratios across Mediterranean-type ecosystems. *Global Ecol Biogeogr* **21**:1147–56.
- Stropp J, Ladle RJ, Malhado ACM, et al. (2016) Mapping ignorance: 300 years of collecting flowering plants in Africa. *Global Ecol Biogeogr* **25**:1085–96.
- Takhtajan A (1986) *The Floristic Regions of the World*. Berkeley: University of California Press.
- Talbot JM, Bruns TD, Taylor JW, et al. (2014) Endemism and functional convergence across the North American soil mycobiome. *Proc Natl Acad Sci USA* **111**:6341–6.
- Tang C, Robson AD, Dilworth MJ (1990) The role of iron in nodulation and nitrogen fixation in *Lupinus angustifolius* L. *New Phytol* **114**:173–82.

- Ter Braak CJ (1986) Canonical correspondence analysis: a new eigen-vector technique for multivariate direct gradient analysis. *Ecology* **67**:1167–79.
- Ter Braak CJF (1987) *CANOCO—A FORTRAN Program for Canonical Community Ordination by Partial Detrended Canonical Correspondence Analysis, Principal Components Analysis and Redundance Analysis (Version 2.1)*. Wageningen: TNO Institute of Applied Computer Science.
- Thuiller W, Araujo MB, Lavorel S (2004) Do we need land-cover data to model species distributions in Europe? *J Biogeogr* **31**:353–61.
- Trinder-Smith T, Cowling R, Linder H (1996) Profiling a besieged flora: endemic and threatened plants of the Cape Peninsula, South Africa. *Biodivers Conserv* **5**:575–89.
- Trinder-Smith T (2006) *Wild flowers of the Table Mountain National Park*. Cape Town: Botanical Society of South Africa.
- Tuohy JM, Prior JA, Stewart GR (1991) Photosynthesis in relation to leaf nitrogen and phosphorus content in Zimbabwean trees. *Oecologia* **88**:378–82.
- Witkowski E, Mitchell D (1987) Variations in soil phosphorus in the fynbos biome, South Africa. *J Ecol* **75**:1159–71.