Supplementary Information S1

Strength, Weaknesses, New Research Questions with Potential, and Key Literature for Biomes and Data Synthesis Methods Discussed in the Text

In this overview there is no relevance to differentiate between uncalibrated (^{14}C y BP) and calibrated years before present (BP) and all ages are expressed as "years BP" (y BP) or "kilo-years BP" (ka). For references to pollen sites the reader is referred to interactive materials in the open access publication by Flantua et al. (2015) and the literature database at http://www.latinamericanpollendb.com/. Other references can be found in the main text.

1. Quaternary Histories of Colombian Biomes

1.1. Savannas of the Llanos Orientales

Strengths

- Nine well–dated pollen record from savanna lakes show a consistent change in environmental conditions from dry (18–11 ka), less dry (11–9.3 ka), again dry (9.3 to 7.1–5.3 ka), increasing precipitation/longer wet season from 7.1–6.4 to 4–3.6 ka, and relatively wet conditions during the last 4 to 3.7 ka.
- The increase in the proportion of palms (*Mauritia* and *Mauritiella*) seems climate driven, but the potential additional driver of human occupation of the savanna has not been established yet.

Weaknesses

- Most pollen sites are located on an east to west transect (following accessibility) hampering analyses of north to south climatic gradients.
- The hypothesis that long–term change in Holocene climatic humidity is driven by the southward migration of the monsoon front intertropical convergence zone (ITCZ) needs supportive evidence.
- There is significant variety between records in the age of climate change events. This seems not to reflect differences in the quality of the age models. Potentially the length of the wet season, i.e. the period of regional rains, extended by the period of effective water availability by stagnant water and wet soils, is driving this variability.

- The start of crop cultivation is unknown; the use of phytoliths and charcoal records may be decisive in identifying the start of human presence in the savanna.
- A robust record of the fire history in the savannas of the Llanos Orientales is lacking and may potentially contribute to clarify the history of human occupation.

Selected Potential Research Questions

- The interpretation of the increase of palms in the period of 4000–3500 BP should be further substantiated by charcoal and phytolith analysis and preferably by an independent proxy of climate change.
- What is the fire history of these savannas and how does it relate to the archaeological record of human occupation?
- Identification of the changing proportions of C3 and C4 plants can be established on the basis of phytoliths and/or molecular biomarkers and identify changes in plant available moisture.
- Is a glacial savanna more rich in C4 grasses compared to today?
- Are C4 grasses running into an endangered status under doubled $p\mathrm{CO_2}$ conditions?
- Are lakes in Casanare and sediment–filled meanders in Arauca able to strengthen the analysis of the north to south gradient?
- Current practice of afforestation of the savanna with rapidly growing exotic tree species (*Eucalyptus, Pinus*) to sequester CO_2 is possibly is not sustainable on a time scale of more than a tree generation.

Selected Literature

Behling & Hooghiemstra (1998, 1999, 2000, 2001), Berrío et al. (2000a, 2000b, 2002a, 2012), Botero (1999), Flantua et al. (2007), Marchant et al. (2006), Mistry (2000), Romero–Ruiz et al. (2012), Rull & Montoya (2014), Sarmiento (1984), Vélez et al. (2005b, 2005c), Wijmstra & van der Hammen (1966), Wille et al. (2003),.

1.2. Amazonian Rainforest

Strength

- Integrated geological and molecular phylogenetic results have strongly driven progress in understanding of the origins of biodiversity in Amazonian rainforest.

Weaknesses

- Available network of pollen sites is insufficient.
- Pollen records mostly come from cores drilled in sediment–filled meanders, and mostly reflect (the younger parts of) the Holocene.
- Pollen morphology of important plant families of the rainforest is insufficiently documented.
- Most pollen taxa in diverse tropical forest occur as singletons challenging the development of adequate main ecological groups in the pollen spectra. This hampers the identification of changing ecological conditions in a rainforest.
- Human impact on Amazonian forest might be older than assumed. Human impact often gives trees with pioneer qualities an advantage over late–successional taxa. However both traits occur as natural elements in virgin forest; local human impact is difficult to substantiate from pollen alone and requires additional records of e.g., charcoal, phytoliths, starch and archaeological records.

Selected Potential Research Questions

- New studies of proxies indicative of human impact (pollen and seeds of crop plants, phytoliths, charcoal, starch, geochemical tracers) are required.
- Modern pollen rain studies in a suite of rainforest environments are required to improve the interpretation of pollen records.
- A study of current practice of agriculture in black soils may provide a reference to better understand "terra pretas" of the past.
- Pollen records from basins beyond the influence of a drainage system are needed to expand our current knowledge into the last glacial, and beyond.
- Marine pollen records from the Amazon fan potentially provide basin–integrated pollen spectra (Hooghiemstra et al., 2006; Hoorn et al., 2017), sufficient to assess changes in large–scale basin configuration, but insufficient to reconstruct regional and local forest histories.

Selected Literature

Behling et al. (1999, 2010), Bermingham et al. (2005), Berrío et al. (2003), Botero (1999), Colinvaux et al. (2000), Goulding et al. (2003), Haffer (1969), Hooghiemstra & van der Hammen (1998), Hooghiemstra et al. (2006), Hoorn & Wesselingh (2010), Hoorn et al. (2010, 2017), Prance (1982), Salamanca–Villegas et al. (2016), Urrego (1994), van der Hammen & Cleef (1992), van der Hammen & Hooghiemstra (2000), Wesselingh et al. (2010).

1.3. Chocó Rainforest

Strenghts

- Chocó rainforest is relatively little disturbed and offers good opportunities to study natural environmental conditions.
- Main drivers of change are extrinsic (sea level change causing migration of mangrove forest and more inland migration of salt tolerant taxa) and intrinsic (changes in the drainage system, flooding frequency, nutrient availability).

Weaknesses

- High pollen and spore diversity is challenging the development of meaningful ecological groups.

Selected Potential Research Questions

- Present–day high sea level stands have minimalized Chocó's surface area. During most of the Pleistocene several offshore islands were connected to present–day mainland Chocó. To which degree sedimentary archives in these islands, and on the floor of shallow offshore waters can contribute to unravel Chocó's history during a glacial–interglacial cycle?
- Pollen records are expected not to be sensitive for changes in precipitation (as current precipitation is by far exceeding the minimally required values), neither in temperature (annual and daily variation is too small), suggesting these forests are well chosen to study intrinsic forest dynamics.
- What is the effect of extremely high precipitation on the altitudinal distribution of plant taxa?

Selected Literature

Behling et al. (1998a), Berrío et al. (2000a), Bush & Colinvaux (1994), Gentry (1982), Vélez et al. (2001)

1.4. Dry Forest of the Inter–Andean Valleys

Strenghts

- There is sufficient basic information to develop a research project with focus on the vegetation dynamics in dry forest since the Last Glacial Maximun.

Weaknesses

- The meteorological constraints in the condensation zone requires a better understanding to improve the understanding the dynamic history of the altitudinal interval with

cloud forest, and to improve the understanding how pollen records are reflecting this dynamics.

- Many dry resistant plant taxa belong to families with insufficient possibilities for pollen identification (Apocynaceae, Bignoniaceae, Bromeliaceae, Burseraceae, Cactaceae, Caparidaceae, Caesalpinaceae, Commelinaceae, Euphorbiaceae, Fabaceae, Malpighiaceae, Malvaceae, Mimosacae, Nyctaginaceae, Poaceae, Setrculiaceae, Rutaceae and Verbenaceae; Ariza, 1999).
- Within the dry forest biome only few swamps in small depressions can be found making sedimentary archives rare.
- No record of a dry forest biome is penetrating into the last glacial; we have no understanding how proportions of C3 and C4 plants have changed altitudinal representation since the Last Glacial Maximun.

Selected Potential Research Questions

- Any new proxy serving a better discrimination between C3 and C4 plants is helpful to detail vegetation dynamics.
- The transition from dry forest in the valley floors and lower slopes to mesic montane vegetation at higher elevations where a temperature–driven condensation zone prevails, needs a better understanding.
- Compared to C3 plants, C4 plants have their optimal profit under conditions of low atmospheric $p\mathrm{CO}_2$ and drought stress in valley floors where rain shadow effects are maximal. Are C4 plants gradually losing their competitive power under increasing atmospheric concentrations of CO_2? What is the future of vegetation rich in C4 plants? (Cavender–Bares et al., 2016).
- Pollen record Pedro Palo suggests that during the Last Glacial Maximun dry forest in the Magdalena Valley was at higher elevations transitional to dry páramo vegetation, without a zone of forest in between. Where were such Last Glacial Maximun scenario's located? Were the Magdalena and Cauca valleys highways for latitudinal migration of dry forest? What is the isolation and connectivity history of dry forest?

Selected Literature

Berrío et al. (2002b), Bullock et al. (1995), Cavender–Bares et al. (2016), Dexter et al. (2015), González–Carranza et al. (2008), Mayle (2006), Pennington et al. (2006), Vélez et al. (2005a).

1.5. Lower Montane Forest (LMF)

Strenghts

- The pollen record of Lusitania shows *Quercus* reached at 1500 m elevation maximum abundance during glacial times when the UFL was close to 2100 m elevation. Pollen record Lusitania shows *Quercus* was an important tree at 1500 m.a.s.l. during the full record of the last ca. 40 ka which supports the understanding that *Quercus* is a component of UMF as well as of LMF.

Weaknesses

- The Last Glacial Maximun position of LMF is assumed between ca. 800 and ca. 1400 m elevation (Hooghiemstra & van der Hammen, 2004). However, LMF dynamics is poorly documented in pollen records.
- Environmental constraints of the lowermost and uppermost boundary (ecotone) of the LMF are poorly understood.
- Pollen from important taxa of the LMF (Cleef & Hooghiemstra, 1984) is pollen morphologically poorly or not documented.

Selected Potential Research Questions

- The transition between lower and upper montane forest in terms of species turnover, changing dominance of traits, and climatological constraints is little documented and poorly understood and needs further research.
- The pollen morphology of leading taxa of the lower montane forest needs better keys for identification.
- How robust is current understanding that LMF taxa are susceptible for night frost and night frost frequency is the main constraint for the altitudinal position of this ecotone.
- To which degree are typical LMF taxa bound to lower montane elevations, during events of rapid climate change in particular? How do arboreal taxa with pioneer qualities migrate over the full altitudinal range of montane forest?
- Most of the LMF has been cleared for plantations. What were the characteristic ecosystem services of LMF to pre–Colombian people?
- In which altitudinal interval is the highest density of archaeological sites in Colombia and how is this interval related to ecosystem services, and availability of suitable living areas?

Selected Literature

Bush et al. (2006), Bush & Hooghiemstra (2005), Cleef & Hooghiemstra (1984), Gentry (1982, 1995), Hooghiemstra & van der Hammen (1993, 2004), Monsalve (1985), Wille et al. (2000, 2001).

1.6. Upper Montane Forest (UMF)

Strenghts

- Most pollen records originate from the UMF belt (Flantua et al., 2015), allowing to develop the most detailed studies in space and time.
- Modern vegetation ecology is a highly advanced discipline, serving optimally the interpretation of pollen records. Understanding of past spatial changes driven by climate change has improved (Flantua et al., 2015, 2016a).
- Long multiproxy records (pollen, TOC, GSDs, XRF–based geochemistry) have generated much understanding how the biotic and abiotic components of the environment have responded to climate change.

Weaknesses

- Understanding of the relationship between modern pollen rain and vegetation is not systematically explored.
- The lowermost UFL position during the Last Glacial Maximun has been indirectly established from pollen record Lake Pedro Palo located at 2000 m.a.s.l. However, this site suggests an uncommon setting of lateglacial vegetation distribution: dry inter–Andean forest was transitional to dry páramo and evidence of a zone with montane forest in between is lacking. It is unclear to which degree montane forest was poorly developed to absent in the driest parts of the northern Andes during dry glacial conditions.
- There is some confusion how rates of change (RoC) should be interpreted as long as a pollen site is immersed in UMF, rate of taxonomic change is low. When an ecotone (e.g., LMF–UMF ecotone) is passing the coring site, the rate of taxonomic change increases. UMF biome–specific RoC (observations from inside the UMF) is expected to be low, supported by the results of Felde et al. (2016). *Site–specific RoC* is expected to show maximal change when the LMF–UMF ecotone, or the UMF–páramo ecotone is crossing the elevational level of the coring site. Biome–specific RoC is poorly understood and subject to speculation. The RoC is influenced by the lengths of the time intervals between the pollen samples (Flantua et al., 2014) which potentially makes comparisons between records difficult.

Selected Potential Research Questions

- Relationships between modern pollen rain and modern vegetation needs a better exploration. Given the high levels of deforestation and degradation of remnants of forest patches at UMF elevations results should be considered with care.
- The relationship between the altitudinal position of the UFL and the in situ AP % needs a systematic assessment to better understand the role of various climate variables, and finally to improve AP %–based temperature reconstructions. Currently establishing the UFL position has an estimated maximum uncertainty of ±100 m which equals to ±1.5 °C (Groot et al., 2013; Hooghiemstra et al., 2012).
- The physiognomy of the UFL needs further study: is the natural UFL reflecting a smooth gradient, or an ecotone with patches of forest in the páramo and patches of páramo in the forest? (Bakker et al., 2008).
- Relevant arboreal taxa making up the UMF need a systematic pollen morphological study and documentation.
- Once raw counts of pollen diagrams can be compared to a database of modern pollen rain and vegetation cover the nearest modern vegetation communities of former non–analogue vegetation associations can be found. In this way changing past forest associations can potentially be given additional interpretation.
- Temperature reconstructions from independent molecular geochemical proxy evidence is needed to verify pollen–based reconstructions.
- A comparison of the last ten interglacials by numerical data analysis showed little taxonomic differentiation between interglacials (Felde et al., 2016). However, the sequence of arboreal taxa expanding their elevational range in each of the interglacials need further analysis. What is the specific role of taxa with pioneer qualities, and what are the lag times of the late–successional taxa? Such understanding from the past potentially improves our ability to better understand current vegetation change and migration as a result of anthropogenic warming.
- Events of rapid climate change allow taxa with pioneers qualities to move altitudinally fast, changing the floral composition of the vegetation in adjacent altitudinal intervals. The contribution of extrinsic drivers (climate) and intrinsic processes, such as ecological competition and legacy effects, are insufficiently understood.
- Long continental records of environmental and climate change are crucial to understand the dynamic nature of ecosystems in the past, what are the boundary conditions that led to present–day ecosystems (see the Funza and Fúquene records), and how ecosystems may respond to future cli-

mate change and anthropogenic impact on the earth system. Although collecting deep cores is facilitated by international drilling programs, accommodating sample analysis with high resolution is a major bottle neck. Formation of consortia of laboratories may be part of a solution.

- Analysis how areas of UMF have been isolated and connected by upper forest line shifts during a full glacial–interglacial cycle will provide crucial understanding which areas can be considered as "stable" or "dynamic" with potential relationships to distribution patterns of biodiversity.

Selected Literature

Bakker et al. (2008), Bogotá et al. (2011a, 2011b, 2015), Bush & Hooghiemstra (2005), Bush et al. (2006), Crawford (2008), Cuesta et al. (2017), Felde et al. (2016), Gómez et al., (2007), González–Carranza et al. (2012), Groot et al. (2011, 2013, 2014), Helmens & Kuhry (1986), Hessler et al. (2009), Hooghiemstra (1984, 2006), Hooghiemstra & Berrío (2007), Hooghiemstra & Cleef (1995), Hooghiemstra & Ran (1994), Hooghiemstra & van der Hammen (2004), Hooghiemstra et al. (2006), Hooghiemstra et al. (2014), Jackson et al. (1996), Kuhry (1988a, 1988b), Marchant & Hooghiemstra (2004), Mora et al. (2010), Moscol–Olivera & Cleef (2009a), Moscol–Olivera & Hooghiemstra (2010), Moscol–Olivera et al. (2009), Muñoz et al. (2017), Sarmiento et al. (2008), Torres, (2006), Urrego et al. (2016), van Boxel et al. (2014), van der Hammen (1974), van der Hammen & Hooghiemstra (2003), Van 't Veer & Hooghiemstra (2000), van 't Veer et al. (1995), Velásquez–Ruiz (2004), Velásquez–Ruiz & Hooghiemstra (2013); Vélez et al. (2003, 2006), Vriend et al. (2012), Weng et al. (2006, 2007), Wille et al. (2002), Wijninga (1996a, 1996b, 1996c, 1996d), Wijninga et al. (2004).

1.7. Páramo

Strenghts

- The floral composition and ecological ranges of páramo taxa are well documented.
- The pollen reference collection of Colombian páramo taxa is well stocked at the University of Amsterdam (but not digitalized yet).
- There are many pollen records available from páramo sites allowing new research based on the LAPD database.
- The availability of long Pleistocene records of páramo dynamics may fuel phylogenetic studies of genera with a strong Pleistocene speciation history.

Weaknesses

- Climatological (and geological?) drivers of the ecotones within the páramo biome are insufficiently understood.
- There is a hint to changing proportions of C4 vs. C3 plants in the páramo at the last glacial to Holocene transition when pCO_2 changed from ca 180 to ca 280 ppmV. However, evidence is limited and implications are poorly understood.
- Estimating the altitudinal position of the UFL from a pollen record located in the páramo is challenging. Wind–blown arboreal pollen cause an overrepresentation of forest pollen, making the rule–of–thumb (UFL corresponds to ca. 40 % AP) invalid. Molecular biomarkers (without wind–blown effects) evidenced this discrepancy (Jansen et al., 2013). The implication is that assessment of the Lateglacial to Holocene transition based on biostratigraphy (in absence of ^{14}C ages) is an unreliable approach.

Selected Potential Research Questions

- Drivers of the ecotones between the three types of páramo vegetation should be better understood, as well as the causes of deviating vertical proportions of the three páramo zones (van der Hammen, 1981, 1985).
- Reconstructing the UFL position from a pollen record located at high elevations in the páramo might be challenging and needs further study, a.o. by supporting analysis of modern pollen rain studies along elevational gradients.
- In the uppermost grasspáramo and lower superpáramo species are near the end of their ability to respond by altitudinal migration to climate warming. How are patterns of plant diversity changing there?

Selected Literature

Boom et al. (2001, 2002), Bosman et al. (1994), Castaño–Uribe (2002), Cleef (1979, 1981), Cleef et al. (1993), Crawford (2008), Ehlers & Gibbard (2004), Guhl (1982), Helmens (2004, 2011), Hooghiemstra et al. (2012), Jansen et al. (2013), Luteyn (1999), Melief (1985), Muñoz et al. (2017), Salgado–Labouriau (1979); Salomons (1986); Sklenář et al. (2005), van der Hammen (1981, 1995), van der Hammen & González (1965b), van der Hammen et al. (1980), Vareschi (1970), Velásquez–Ruiz & Hooghiemstra. (2013).

2. Multi–Site Syntheses

2.1. Latin American Pollen Database (LAPD)

Strenghts

- The number of available pollen records in the LAPD, the quality of the age models, and the temporal resolution is remarkably high compared to other continents (Flantua et al., 2015). This allows to use the LAPD for new research questions.
- Results of the 2014–update of the LAPD have been published open access.

Weaknesses

- The metadata of many pollen sites are still not supported by the original raw counting data. As a result, synthesis work in Latin America is hindered.
- In absence of a data base manager, continuation of the LAPD depends on the willingness of individual authors to contribute with the metadata and raw counting data of new sites.

Selected Potential Research Questions

- The community of paleoecologists is increasingly ready to submit raw data to public domain data repositories such as Neotoma. There is an urgent need to organize data management, preferable as a collaborative effort of a consortium of researchers.
- Biomisation of LAPD data using signals of crop cultivation and degraded vegetation as an "artificial plant functional type" (PFT) in combination with an improved fire record, allows new links with archaeology (Marchant et al., 2001a, 2001b).

Selected literature

Flantua et al. (2015, 2016a, 2016b), Rull & Montoya (2014), Smith & Mayle (2017).

2.2. Paleodata–Model Comparisonas: The Biomisation Method

Strenghts

- Results are replicable and quantitative.
- For most of the pollen taxa in pollen diagrams from the Neotropics the habitat and ecological range is known (Marchant et al., 2002c). Most pollen types recognized in Colombian pollen records have been assigned to PFTs, and all PFTs have been assigned to biomes.
- A reconstruction of modern biomes shows that, for the majority of the pollen sites, the pollen data accurately reflect the potential vegetation, even though much of the original vegetation has been transformed into agriculturally used area.
- The biomisation method has been designed for spatial analysis. However, each pollen sample from a pollen record can be biomised (downcore biomisation) leading to a record showing in time how traits (close equivalents of PTFs) are changing of proportions and reflecting how vegetation is responding physiologically to climate change (Crawford, 2008; Marchant et al., 2002c; Woodward & Cramer, 1996).
- The number of PFTs and biomes is not restricted, thus allowing to design the biomisation method for special research objectives.
- Grouping pollen types of crop plants, pioneer taxa after disturbance, and taxa indicative of degraded vegetation in a PFT, human impact on the vegetation can be shown in space and time (Marchant et al., 2001b). This method allows to develop new links between paleoecology and archaeology.

Weaknesses

- The biome with an assigned value below the maximum value is discarded. This step in the procedure reflect a loss of information. Some alternatives to avoid information loss have been developed.

Selected Potential Research Questions

- The updated LAPD allows a new and more detailed temporal–spatial assessment of human impact in Colombia.
- Downcore biomisation may support a further understanding how changing forest composition expressed in changing proportions of traits, is reflecting climate change.

Selected Literature

Flantua et al. (2015), Marchant et al. (2001a, 2001b, 2002a, 2002b, 2002c, 2004a, 2004b, 2009), Prentice et al. (1992, 1996).

2.3. Pollen–Driven Distribution Areas in a Digital Elevation Model

Strenghts

- Long pollen records are a rich source of information for evolutionary processes in deep ecological time. New inter-

disciplinary links may be developed between paleoecology, and molecular phylogenies framed in geological time.

- Quantitative spatial and temporal pattern analysis allows to calculate "projections", or even to generate higher confidence "predictions" for the duration of connectivity and isolation of plant communities (the gene pools for evolutionary processes).

Weaknesses

- Estimates of the UFL position are with an estimated ±100 m uncertainty envelope, reflecting 10 % to 15 % of the total UFL amplitude during a glacial–interglacial cycle of maximally 1500 m. For specific locations UFL shifts may differ from the values in inferred from the Funza and Fúquene records, depending local climatological conditions (see Marchant et al., 2001a). Remarkable enough, current estimates of sea level change (both sea level change and UFL shifts are dependent on mean global temperature) reflect a similar level of uncertainty (Spratt & Lisiecki, 2016).

Selected Potential Research Questions

- Connectivity and isolation of plant populations (gene pools) during most of the Quaternary can be quantified in length of time (ky) and altitudinal interval (m.a.s.l.). However, current distribution patterns reflect the "end stage" of a long process of expansion and loss of distribution area. Therefore, a simple relationship between past frequency of isolation and connectivity with current patterns of diversity is not expected.
- Various capacities among seeds to be transported by birds, and various pollination syndromes potentially influence the efficiency of geographical isolation. Mapping vegetation across cordilleras with stepping stones on the basis of dispersal capacities may contribute to understand past migration routes.

Selected Literature

Flantua & Hooghiemstra (2017, 2018), Flantua et al. (2014).

3. Human Impact on the Environment

Strenghts

- A growing number of pollen records show around 3 ka a distinct expansion of human impact on the environment. Locally, human occupation histories are starting earlier or later.

Weaknesses

- The pollen morphology of many crop plants is insufficiently known. Analysis of paleobotanical remains (seeds, etc.), phytolith, starch grains, and molecular biomarkers may provide important additional evidence (Jansen et al., 2013; Piperno, 2006).

Selected Potential Research Questions

- Re–analysis of the timing and distribution of human impact on the basis of an artificial PFT (see Marchant et al., 2001b) is a challenging exercise to build a stronger link between paleoecology and archaeology.
- A pollen morphological and phytolith study of crop plants is very welcome.
- "Natural vegetation" is often considered as the opposite of "disturbed vegetation". As "natural vegetation" bears the unknown legacy of many disturbing events in the past, many different versions of a "natural vegetation composition" are expected. However, during the last million years, ten disturbance events of glacial conditions have impacted the composition of north Andean interglacial forest so little that no single interglacial forest is characteristic enough to be recognized individually (Felde et al., 2016). External (extrinsic) and internal (intrinsic) environmental processes and constraints leading to "stable" and "dynamic" forest communities, on time scales of 10^2 to 10^4 years, are little understood.

Selected Literature

Botiva–Contreras et al. (1989), Castaño–Uribe (2002), Flantua et al. (2016a), Gómez et al. (2007), Scarre (2013), Schreve–Brinkman (1978), van der Hammen & Correal–Urrego (1978).