Understanding controls on stanols in lake sediments as proxies for palaeopopulations in Mesoamerica

Benjamin Keenan[1], Anic Imfeld[2], Yves Gélinas[2], and Peter M. J. Douglas[1]

[1]Department of Earth and Planetary Sciences, GEOTOP-McGill University, Montréal, QC H3A 0E8, Canada

[2]Department of Chemistry & Biochemistry, Concordia University, Montréal, QC H4B 2A7, Canada

Corresponding author: Benjamin Keenan (benjamin.keenan@mail.mcgill.ca)

Faecal stanols; Coprostanol; Palaeo-population; Archaeological demography; Central America; Groundtruthing

Faecal stanols in lake sediments have been used as a proxy for human populations in the past in a variety of contexts, with the assumption that variability in faecal stanol concentration or ratios is a reliable proxy for relative catchment-scale human populations. Despite that, the specific controls on faecal stanol concentrations and ratios in lake sediments remain poorly understood. In this study we analyse faecal stanol concentrations in lake surface sediments across Guatemala and the Yucatán Peninsula of Mexico in order to constrain geographical and biogeochemical variables controlling stanol concentrations and ratios in lake sediments in this region. We propose and test the hypothesis that the stanol ratios coprostanol:(coprostanol+stigmastanol) and coprostanol:(coprostanol+cholestanol) scale according to the proximity to and size of nearby population centres. The key controls on stanol concentrations that we identify are the proximity to human population centres and the human population within 5 km of the sampling point. The relationship between coprostanol and stigmastanol suggests a human origin for stigmastanol at Lake Petén Itzá, which has a much larger human population in its catchment, but an herbivore origin at other lakes. Based on a transect across Lake Petén Itzá, the ratio coprostanol:(coprostanol+cholestanol) does not appear to be an accurate proxy for proximity to human population centres, nor does it correlate with human population. We suggest that normalising stanol concentrations to TOC is an appropriate way to take into account the effects of mineral dilution as well as the potential effects of organic matter deposition and preservation, and that the ratio coprostanol:(coprostanol+stigmastanol) may be an effective approach to determine the relative contribution of coprostanol-producing mammals and herbivores, but does not scale with human population. Further, we discuss the current limitations of the proxy as well as its future directions, including the implications of our results for sediment core siting, the use of stanol ratios in palaeolimnology, as well as the storage, transport, and diagenesis of stanols.

Introduction

Faecal stanols are lipid biomarkers produced in animal intestinal tracts and their distribution is determined by diet, the species of animal, and the presence and taxonomy of anaerobic bacteria in their digestive tracts (Bethell et al. 1994). Coprostanol (5β-cholestan-3β-ol) is produced during metabolic reduction of cholesterol in the intestinal tract of most mammals and is the major sterol present in human faeces (Bethell et al. 1994; Bull et al. 1999). Of the mammals capable of producing coprostanol, pigs, sheep and cows are known to produce sufficient concentrations to potentially mask the signal of human coprostanol (Prost et al. 2017; Zocatelli et al. 2017). The identification and quantification of coprostanol has been used as a means of determining anthropogenic impact by determining the extent of sewage discharge in coastal sediments (Hatcher and McGillivary 1979; LeBlanc et al. 1992; Hussain et al. 2010) as well as archaeological and palaeoenvironmental studies (Bethell et al. 1994; Bull et al. 2003; D’Anjou et al. 2012; Sistiaga et al. 2014; White et al. 2018; Vachula et al. 2019; Sear et al. 2020). In some cases faecal stanols have been quantified alongside other markers such as polycyclic aromatic hydrocarbons (PAHs) in order to characterise relationships between industrial activity, vehicle traffic and hormones (Lima et al. 2019). Faecal stanols have also been quantified in coprolites and used as a proxy for diet (Zhang et al. 2020). Stanol analysis of fish muscle has given insights into the dietary changes of detritivorous fish reflecting anthropogenic organic matter close to sewage dominated inputs (Speranza et al.

2020). Stanols have been used to trace marine mammals and birds in marine sediments (Venkatesan et al. 1986; Venkatesan and Santiago 1989).

After production, stanols can enter the environment as a component of faeces particles. These lipid biomarkers are relatively resistant to degradation and accumulate in sediments within depositional settings such as marine and lacustrine basins (D'Anjou et al. 2012; Pratt et al. 2008). The conversion of coprostanol to epicoprostanol (5β-cholestan-3α-ol), mediated by microbes, can occur *in situ* in both soils and sediments (Bull et al. 2002). The relative proportions of 5β-stanols can also be used as a means of determining the input of faecal matter from animals with different diets (Leeming et al. 1996; Sistiaga et al. 2014). For example, faeces of herbivores dominantly contain 5β-stigmastanol, followed by β-sitosterol or 5α-stigmastanol, because of the prevalence of plants in their diets (Prost et al. 2017). Therefore analysis of 5β-stigmastanol can help to control for contributions of coprostanol from grazing herbivores, which produce coprostanol in lower proportions than humans.

Further work has applied various diagnostic ratios in order to determine the source of faecal waste. In the application to lake core records, faecal stanols have been discussed as absolute concentrations in dry sediment, as ratios to other stanols and sterols, and as ratios to total organic carbon (TOC). Stigmastanol, stigmasterol, sitosterol, cholestanol, and cholesterol have also been characterised (D'Anjou et al. 2012; Battistel et al. 2017; Prost et al. 2017; White et al. 2018; Vachula et al. 2019; Shillito et al. 2020; Keenan et al. 2021).

The most appropriate approach to use in reconstructing human populations, as well as the environmental controls on faecal stanol concentrations and ratios in lake sediments, remain poorly constrained. Faecal stanols in tropical lake sediments, and

Mesoamerican lake sediments in particular, are understudied, despite their potential to complement archaeological estimates of demographic change in this region (Escobar et al. 2020, Keenan et al. 2021). This novel proxy needs to be further refined in order to accurately reconstruct past human population change. To accomplish this we analyse faecal stanols in a set of modern lake sediments from Petén, Guatemala, and Yucatán and Quintana Roo, Mexico, in order to better understand the controls on stanol concentrations and ratios in tropical lake sediments.

Lake surface sediments collected from 9 lakes (Fig. 1) across the Yucatán Peninsula and northern Guatemala were analysed in order to understand the geographic and biogeochemical variables that control faecal stanol concentrations and ratios. The lakes were chosen in order to analyse stanol concentrations and ratios across lake catchments that vary in hydrogeochemistry and human population. Specifically, we tested whether coprostanol concentrations and ratios are higher in lakes closer to population centres and whether limnological properties that potentially influence stanol preservation are correlated with coprostanol concentrations. We examined intra-lake variability in coprostanol concentrations and ratios within one large lake basin, namely Lake Petén Itzá. Further, We tested the hypothesis that the stanol ratios coprostanol:(coprostanol+stigmastanol) and coprostanol:(coprostanol+cholestanol) scale according to the proximity to and size of nearby population centres.

Study site

Lake surface sediments were collected from 3 lakes in the Mexican lowlands (Yucatán Peninsula) and 6 lakes in the Guatemalan lowlands (Petén department of northern Guatemala). These lakes are found on the extensive karst area of the Yucatán

Platform Province where groundwater moves through submerged cave systems produced from the dissolution of carbonate rock (Pérez et al., 2011). Lakes in the Petén Department (Perdida, Sacpuy, Petén Itzá, Macanché and Salpetén) are also part of this limestone platform. All of the lakes with the exception of Noh Bec have stratification. Stratification data for Lake Cobá is unavailable (Pérez et al., 2011). The "Hotel" lake has not previously been studied. A table of information including depths and surface areas for the lakes is available in the Electronic Supplementary Material (*ESM1*).

Materials and methods

Lake surface sediment sampling

Lake surface sediments from the Yucatán Peninsula and northern Guatemala were collected using an AMS Ekman Dredge in July of 2019, sampling approximately the uppermost 3 cm of sediment. The mud-water interface was retained. The sampling locations were determined based on literature to encompass a range of biogeochemical variables, including salinity, as well as satellite imagery, and site access. They represent a range of water depths, lake properties, land-use, and proximity to human population centres (Table I). Sediments were stored in a cool box in the field, transferred to a refrigerator, and stored at 2 ºC for up to two weeks in Montréal before being frozen and freeze-dried prior to analysis to remove water. Water chemical property measurements on a depth transect from 0 (surface) to 10 m water depth were taken where possible using a YSI professional plus multiparameter instrument configured to record dissolved oxygen, pH, resistivity, specific conductivity, and TDS.

Stanol quantification

Dried sediment samples were ground, weighed, added to a PTFE tube and extracted using a CEM MARS 6 microwave extractor with 10 mL of 9:1 dichloromethane:methanol. This ratio of solvents was selected after testing various methods for their extraction efficiencies using lake sediment samples (Kornilova and Rosell-Melé 2003; Battistel et al. 2015). The MARS 6 oven was heated to 80 ^{o}C and held at that temperature for 20 minutes. The contents of the PTFE tube were then transferred to a centrifuge vial, centrifuged and the Total Lipid Extract (TLE) was transferred to an evaporating vial. A few mL of 9:1 dichloromethane:methanol were added twice more to the centrifuge tube in order to ensure complete removal of extracted material. The TLE was evaporated and split into 2 fractions (a non-polar fraction and a polar fraction) using silica gel chromatography. The pipette columns consisted of 5 cm of silica gel, and 1 cm of sodium sulphate. 15 mL of hexane was eluted to collect the non-polar hydrocarbon fraction and 15 mL of methanol was eluted to collect the remaining neutral and polar fractions. The polar fraction was saponified using KOH (potassium hydroxide) and separated into a neutral and acid fraction. Following 3 rounds of liquid-liquid extraction, the sterol fraction was then derivatised with BSTFA (bis-trimethylsilyltrifluoroacetamide).

The neutral (sterol) fraction was analysed using gas chromatography with a flame ionisation detector (GC-FID) with a TRACE TR-5 GC Column (60 m × 0.25 mm) at McGill University in sequence with known standards for coprostanol, epicoprostanol, and stigmastanol, cholestanol and cholesterol (Sigma-Aldrich) in order to quantify these compounds. A set of representative samples were analysed using an Agilent 7890B GC with an Agilent 5977B MSD (DB-5MS 25 m × 200 µm × 0.33 µm) at Concordia University to confirm compound identification by comparing ion fragmentation with NIST library. Because of the overlapping retention time of coprostanol and epicoprostanol it was not possible to consistently resolve these molecules. We followed the approach of White et al. (2018) and reported the sum of the two. This does not influence our interpretations since epicoprostanol is a transformation product of coprostanol, and therefore the summed concentration represents the net input of coprostanol to lake sediments.

Total Organic Carbon

To measure total organic carbon concentration (TOC), dried and ground sediment samples were first weighed into an open silver cup, placed into a clean tray and fumigated in a closed glass container with a concentrated HCl for 24 hours to remove inorganic carbon. The silver cup was then sealed with tweezers and placed in a tin cup, which is a better catalyst for flash combustion analysis, and analysed with a Carlo Erba NC2500 elemental analyser (Hélie, 2009). These analyses were performed at the GEOTOP Light Stable Isotope Laboratory at the Université du Québec à Montréal.

Population data and sewage management

Population estimates were obtained for populated areas within 5 km of the sampling points at each lake. A 5 km radius was selected as this encompasses at least one population centre for the majority of the sampled modern lakes, with the exception of the sampling point at Lake Chichancanab, which falls out of this range. Furthermore, our data from Lake Petén Itzá implies that coprostanol concentrations decrease substantially within 3 km of the major population centre. Population figures were obtained by analysing the Humanitarian Open StreetMap Team Open Street Map (HOTOSM) in QGIS, an open source desktop geographic information system application. The OpenStreetMap project combines crowdsourcing and field surveying to obtain accurate population data at the local level (Humanitarian OpenStreetMap 2020). Each settlement has a population number associated with it, regardless of the size. The smallest settlement is the village “Hacienda Laguna Perdida”, which has a population of 684.

In San Benito, adjacent to Santa Elena, there is a sewage system and treatment plant to the west of Santa Elena and Flores with a lagoon system built in 2005. However, owing to collapses in levee slopes of two lagoons, the plant was operating at a reduced capacity (Rodas, 2011). Phase 2 of the treatment plant involved the installation of an effluent treatment plant from the discharges of the municipalities of San Benito and Flores and was inaugurated in 2016 (Ministerio de Ambiente y Recursos Naturales, 2016). For the year 2017 the urban area of the municipalities of Flores and San Benito do not have a wastewater management plan, causing wastewater to continuously discharge pollutants in the Lake Petén Itzá basin (Constanza, 2018). The other lakes sampled are surrounded by much smaller towns or villages, and are likely to use septic tanks with drainage fields (David Kuhn, personal correspondence).

Results

Range of variation of key measurements

There is considerable variability in the concentrations of all faecal stanols across all lakes, as well as within lakes with multiple samples. We present statistical measurements of all lakes, lakes excluding Petén Itzá, and Petén Itzá only, given that 15 samples were collected along a transect in Lake Petén Itzá. There is a major human population centre, Santa Elena and surrounding cities, with a population of approximately 70,000 inhabitants, on the shoreline of Petén Itzá. The total mass normalised concentrations of coprostanol+epicoprostanol (from here on, "coprostanol") range from 3 ng/g to 3410 ng/g. The average for all lakes is 570±780 ng/g (± 1σ standard deviation). For all lakes excluding Petén Itzá the average is 390±500 ng/g, and for Petén Itzá alone the average is 790±980 ng/g. The highest absolute concentrations of coprostanol were observed at the sampling location closest to Santa Elena in Lake Petén Itzá (3410 ng/g) and Lake Chichancanab (1820 ng/g).

For stigmastanol average concentrations were 2320 ±5870 ng/g, 3500±8180 ng/g, and 1220±1190 ng/g, for all lakes, all lakes excluding Petén Itzá, and only Petén Itzá, respectively. The highest concentration of stigmastanol is found in Lake Chichancanab (33700 ng/g). Stigmastanol is generally present in much higher concentrations than coprostanol in these lakes, with the exception of Petén Itzá. For cholestanol average concentrations were 880±1390 ng/g, 630±1140 ng/g, and 1200 ±1620 ng/g, for all lakes, all lakes excluding Petén Itzá, and only Petén Itzá, respectively. For cholesterol average concentrations were 970±1380 ng/g, 960±1710 ng/g, and 1050±1020 ng/g, for all lakes, all lakes excluding Petén Itzá, and only Petén Itzá, respectively.

Total Organic Carbon

The average TOC (g/g dry sediment) is 32±15 $\times 10^{-2}$ g/g, 37±16 $\times 10^{-2}$ g/g, 30±1 $\times 10^{-2}$ g/g, for all lakes, all lakes excluding Petén Itzá, and only Petén Itzá, respectively. The average coprostanol normalised to TOC is 1870±2670 ng/g OC, 1480±2760 ng/g OC, and 2410±2650 ng/g OC, for all lakes, all lakes excluding Petén Itzá, and only Petén Itzá. The average stigmastanol normalised to TOC is 10090±35830 ng/g OC, 16650±50500 ng/g OC, and 3760±3350 ng/g OC, for all lakes, all lakes excluding Petén Itzá, and only Petén Itzá. The average cholestanol normalised to TOC is 3080±5810 ng/g OC, 2710±6750 ng/g OC, and 3690±4970 ng/g OC, for all lakes, all lakes excluding Petén Itzá, and only Petén Itzá. The average cholesterol normalised to TOC is 3700±7570 ng/g OC, 4210±10290 ng/g OC, and 3550±3490 ng/g OC, for all lakes, all lakes excluding Petén Itzá, and only Petén Itzá. Fig. 2 shows the relationships between each stanol and TOC, none of which are significant at $p \leq .05$. All stanols and TOC have weak positive relationships at Petén Itzá and very weak negative relationships for all other lakes.

Linear relationships between stanol concentrations

For all lakes the relationships between all of the stanols are positive and significant at $P < .05$ but coprostanol has a weaker relationship with stigmastanol (Fig. 3). When all sampling points from Petén Itzá are excluded to look at variation between lakes without a major human presence, the relationships between the stanols and sterols are all positive and significant, except for cholesterol *vs*. stigmastanol and cholestanol *vs.* stigmastanol which are not significant at $P < .05$. The Petén Itzá samples alone also indicate significant positive relationships between the stanols. A table of regression statistics is available in the Electronic Supplementary Material (*ESM4).*

Relationship with population

All but one lake (Lake Chichancanab) were within 5 km of at least one population centre. In order to assess the relationship between catchment population and stanol concentrations we focused on the samples closest to the shoreline – at Lake Petén Itzá this includes the sampling point closest to Santa Elena and the sampling point closest to San Jose, the two population centres at either end of the transect (Fig. 1). The samples collected along the rest of the transect are not included in this analysis because the population data are not of sufficient resolution to determine which population centres are contributing to concentrations at those sampling points. A spatial relationship between coprostanol and population is not evident at the other lakes sampled, partly because the sampling strategy did not involve sampling along a transect at those lakes.

Coprostanol and stigmastanol for lakes excluding Petén Itzá have positive correlations with population ($r = 0.74$ and 0.74 respectively). When the samples from Petén Itzá are included the r-value for coprostanol increases ($r = 0.97$) but for stigmastanol decreases ($r = 0.38$). Cholestanol and cholesterol are also positively correlated with population, for all lakes. There is no apparent relationship between population and the stanol ratios coprostanol:(coprostanol+stigmastanol) and coprostanol:(coprostanol+cholestanol) (Fig. 4e,f).

Spatial variability in stanol concentrations in Lake Petén Itzá

Absolute concentrations of coprostanol, stigmastanol, cholestanol and cholesterol, and the concentrations of these stanols normalised to organic carbon all decrease rapidly within the first 500 m from the population centre of Santa Elena (Fig. 5). There is then a more gradual decrease until 2200 m from Santa Elena. Absolute concentrations remain relatively low across the rest of the transect. An exponential relationship between concentrations and distance provides a partial fit to these data, but under predicts concentrations near Santa Elena. Coprostanol concentrations normalised to OC show a similar pattern, but increase slightly on the northern side lake close to the population centres of San Andrés and San José on the northern side of Lake Petén Itzá. Coprostanol:(coprostanol+cholestanol) has been used previously (White et al. 2018) as has coprostanol:stigmastanol (Vachula et al. 2019) and cop:(cop+stigmastanol) (Prost et al. 2017). Fig. 5e shows that coprostanol:(coprostanol+cholestanol) does not have a strong correlation with distance. Cop:(cop+stigmastanol) (Fig. 5f) however is relatively high close to Santa Elena, decreases with distance from it, and increases again closer to the population centres on the northern side of the transect. The ratio is higher around 6.5 km from Santa Elena even though the population centres on that end of the transect (San Andrés and San José) have a smaller population.

Discussion

Spatial variability of stanol concentrations within lakes

Stanol concentrations and ratios along the Petén Itzá transect

Lake Petén Itzá has a higher coprostanol concentration in near shore sediments than all the other sampled lakes, consistent with its much larger local population (Fig. 5). Unnormalised concentrations and concentrations normalised to TOC (*ESM2*) decrease with distance from population centre. The high concentrations of stanols close to Santa Elena in Lake Petén Itzá, and the approximately exponential decrease in concentrations with distance from the population centre, suggest that distance from sources of faecal matter (i.e., human population centres) is a key variable controlling concentrations of faecal stanols, at least within a large lake basin. Faecal stanols have been shown to have low solubility in water and associate with particulate matter (Lloyd et al. 2012). Our data shows this to be true in a lacustrine basin, and that coprostanol, and other stanols, are likely transported as a solid particle either in suspension or in remobilised sediments. Further, in the tropical lake basins studied there is a likelihood of dilution with authigenic sediment (i.e. carbonates).

Over 2 km the concentrations of all stanols decrease at a rate that approximates an exponential decline (Fig. 5a-d), with the largest decrease within the first 500 m. The low concentrations closer to the smaller towns on the northern side of the lake are consistent with coprostanol concentrations scaling with population, although different waste management infrastructure, currents or sediment flow pathways could also influence this difference. The northern part of the lake is also much deeper, in excess of 40 metres water depth, compared with 2 metres in the proximity of Santa Elena. This could be important as greater depths provide more opportunity for dilution and mixing prior to deposition (Reeves and Patton 2005).

The exponential decrease in stanol concentrations is not seen in the ratios of coprostanol:(coprostanol+cholestanol) and coprostanol:(coprostanol+stigmastanol) (Fig.5e,f). Ratio values are low close to Santa Elena and for most of the transect, with two exceptions at around 1 km and 1.5 km from Santa Elena.

Stigmastanol concentrations follow a similar trend to coprostanol concentrations at Petén Itzá (Fig. 5b) that might be explained by the association of livestock associated with the population centre at Santa Elena, a combination of livestock and human-derived stigmastanol, or a mixture of the two. Stigmastanol has been reported as a component of not more than half of human coprostanol faecal content (Prost et al. 2017) and is widely considered to be a marker of herbivore faecal contamination (Prost et al. 2017; Harrault et al. 2019).

Spatial variability of stanols in other lakes

The significant spatial variability of concentrations within sediments at Lakes Chichancanab, Noh Bec and Sacpuy (Table I) could also reflect basin-scale heterogeneity in concentrations, but we did not sample along a transect at these lakes. At Sacpuy this difference may be related to lake depth as well as distance to shore, since the sample closer to the shoreline has a higher concentration of coprostanol. At Noh Bec coprostanol concentration appears to be more closely related to distance from the shoreline, as concentrations decrease from 750 ng/g to 40 ng/g over 167 m, between 405 and 572 m from shore. There is no difference in water depth at these sites. At Chichancanab the sample furthest from the shore (220 m) has the highest coprostanol concentration (1800 ng/g) of any sample excluding those from Petén Itzá. This could reflect a unique depositional setting or source of coprostanol in Lake Chichancanab or a potential effect of high salinity (*ESM3*).

Implications for core siting

The significant spatial variability observed in this study in the lakes with multiple samples suggests that future targeting for cores with the intention of reconstructing population using stanols in large lakes should take place as close to the population centre/site of interest as possible, whilst remaining a depocentre. Sediment must still be able to accumulate without risk of flushing by storms, for example. There are few studies looking at the spatial variability of lipid biomarkers, but substantial heterogeneity on a basin-scale has been reported (Sarkar et al. 2014). Faecal stanols have been shown to associate with particulate matter and sediment out quickly (Lloyd et al. 2012), and, unsurprisingly, the greatest concentrations faecal stanols have been found where point sources of sewage discharge have historically been the highest (Hatcher and McGillivary 1979; Murphy et al. 2016). Clearly whether waste input to a lake comes from a point source—i.e., a town or city versus a population diffused across a landscape—is also important for selecting a coring location. In addition, in smaller lake basins the depocentre may be sufficiently close to the population centre to avoid a substantial dilution of the coprostanol signature, but must still be a depocentre.

Relationship of stanol concentrations to local population

The quantification of faecal stanols in modern lake sediments reveals a number of important insights that are relevant to their use as an emerging proxy for population change in the past. Fig. 4 shows that within this dataset concentrations of all stanols are sensitive to population within a 5 km radius, in that they are all positively correlated with population. When samples from Lake Petén Itzá are included, we observe stronger positive relationships with all stanols and sterols except for stigmastanol and cholesterol (*ESM5*). Therefore we infer that stanol concentrations are influenced by population, but that the strongest effect is for coprostanol and cholestanol, when all lakes are taken into account. The total dataset correlation is highly influenced by one point from Petén Itzá, resulting in a slope difference of an order of magnitude. The differences between lakes suggest a non-linear effect, or substantially different slopes between lake environments, suggesting that other variables do play an important role in controlling spatial variations in concentrations. This difference could also point to potential differences in sewage treatments between Petén Itzá and other lakes leading to a lower response per unit of population at Petén Itzá. We also note that based on our results from Lake Petén Itzá, concentrations decline within a short distance of a population centre. Therefore a smaller radius may be more accurate for identifying the relationship between population and coprostanol, but census data are not available at this resolution for Guatemala and Mexico. However Lake Petén Itzá is not representative of most lakes in this environment. The rate of decline in Petén Itzá may be very different than in other lakes. More hydrodynamic environments would lead to higher dispersion and thus lower slopes, for example.

Some sampling sites in this study are close to towns, such as at Cobá (population 1278), but have low concentrations and ratios. In contrast, the Lake

Chichancanab sampling sites were not located near major human settlements, but there are high concentrations of both coprostanol and stigmastanol. We observed that there are people living in the vicinity of Lake Chichancanab, but the census data are not sufficiently spatially resolved in order to account for them. We infer that the variability observed in this sample set is therefore be controlled by factors other than population (*ESM3*).

In the application of the faecal stanol proxy to a lake sediment core the ability to compare patterns in concentrations and ratios with archaeological evidence is limited by the temporal resolution and availability of archaeological material. General patterns emerge in all of the studies discussed here when data are compared with archaeological evidence. In order to refine this proxy further it would be valuable to quantify faecal stanols from a lake sediment core adjacent to a town or city with reliable census data in order to determine the efficacy of the method as a proxy for determining population in the past. When pursuing this approach it is important to bear in mind how waste management strategies might also have changed through time. The sanitation situation in the studied modern lakes is different than in the past.

Use of ratios

The presentation and interpretation of faecal stanol data has been approached in various ways in order to take into account variations in degradation rate as well as low stanol concentrations. For example, White et al. (2018) report their data as a ratio of coprostanol to coprostanol+5α-cholestanol, a stanol commonly found in lake environments. This approach relies on the premise that the ratio of coprostanol to 5α-cholestanol correlates with the amount of human waste transported to the lake, and that the amount of human waste correlates with population size (D’Anjou et al. 2012). In our dataset there is a correlation between coprostanol+epicoprostanol and cholestanol (Fig. 3a), but this ratio does not correlate with population in modern lake sediments. This suggests that although the processes controlling the deposition or preservation of coprostanol and cholestanol are similar, the ratio does not scale with population. In Lake Petén Itzá, the coprostanol:(coprostanol+cholestanol) (Fig. 5e) is low closest to the city and is variable along the transect. This does not support the hypothesised relationship with proximity to population centre.

The controls on concentrations of cholestanol are not well understood, and might limit the use of using the ratio of coprostanol to cholestanol as a means of reconstructing human populations in the past. Because of the relationship of proximity to population centre and coprostanol concentrations, it appears that a ratio of coprostanol to cholestanol is not appropriate. However the failure of our hypothesis that stanol ratios scale according to proximity to and size of nearby population centres does not necessarily preclude the use of these ratios in their use as sewage indicators or as markers of human presence of a landscape or relative population change in the past. Total faecal stanols were used to discuss faecal stanol data from northern Norway (D’Anjou et al. 2012). However, absolute concentrations can be misleading owing to the effects of sediment dilution, contrasting mineral composition, sediment grain size distribution and organic matter content (Bull 2002). It might therefore be more appropriate to normalise absolute concentrations to TOC. Stanol abundance normalised to TOC can be used to take into account the effects of mineral dilution as well as the potential effects of organic matter deposition and preservation on stanol concentrations, thus helping circumvent some of issues that could skew interpretations of stanol data in reconstructing human population, especially in core records covering long time periods (LeBlanc et al. 1992; Thienemann et al. 2017). This does not prevent dilution with organic sediment, which is presumably what we observe with distance from shore in the Petén Itzá transect.

Fig. 2 shows that TOC has no significant relationship with each stanol, for lakes excluding Petén Itzá, as well as for Lake Petén Itzá. This shows that the processes controlling the deposition or preservation of TOC are dissimilar to those for coprostanol and cholestanol, and other stanols. In modern samples TOC is primarily controlled by production and deposition, whereas in sediment cores preservation would be of greater importance. Using TOC as an independent variable to normalise coprostanol to reduces the variation associated with normalising to cholestanol, where the controls on cholestanol concentrations could be related to either human or non-human sources, as well as the advantages described above. Stanols normalised to TOC in the Lake Petén Itzá transect have a similar pattern of exponential decrease that absolute concentrations have, showing that at least in Lake Petén Itzá the normalisation does not skew the patterns to the extent that the data become useless.

In order to assess the relative contribution of humans (and other coprostanol producing mammals) versus herbivores a ratio of coprostanol to stigmastanol can be used. Vachula et al. (2019) use a value of >0.18 for the coprostanol:stigmastanol to infer human faecal contribution, based on Beringian megafauna. Although in our data the ratio does not correlate with population, if the coprostanol:(coprostanol+stigmastanol) ratio is used (Fig. 5f), the transect ratio shows a pattern that might be expected if inefficient sewage treatment existed in Santa Elena, and no sewage treatment existed on the northern side of the transect at the population centres of San Andrés and San José. Prost et al. (2017) use the same ratio with an upper threshold for human presence (0.72) and a threshold for herbivores (0.29), to assess the relative contribution of humans and herbivores to lake sediments. Although there is only one data point along the transect above 0.72, the data points closest to the city are greater than 0.29. If there are additional herbivore contributors (i.e., non-human) then the low ratio could be the result partly of waste processing, as coprostanol input would decrease but stigmastanol would be unchanged. If humans are responsible for all stanol input then we would not expect the ratio to change, as both coprostanol and stigmastanol would be affected.

In situ production of stanols in sediments

Cholestanol and population have a moderate positive correlation, which, like the patterns seen in the Lake Petén Itzá transect where all stanols, including cholesterol and cholestanol decrease rapidly within the first 500 m, suggest that the source of cholestanol might be human-derived or related to human activities, or that microbes convert human-derived cholesterol into cholestanol, or both (Teshima & Kanazawa, 1978). This could take place through delivery of cholesterol and cholestanol via faecal waste from the town, which has poor waste management facilities, or through run off from animal waste used as fertiliser.

Cholestanol (5α-cholestanol) is produced from microbial degradation of cholesterol in the environments (Leeming et al. 1996; Prost et al. 2017). It is also possible that human-influenced processes in soils, such as nutrient inputs via fertiliser, enhance cholestanol production. The relationship between cholestanol normalised to TOC and population is weaker, but still positive. In White et al. (2018) a ratio of coprostanol:(coprostanol+cholestanol) was used to reconstruct population change at Cahokia, Illinois, and compares well when with archaeological evidence for population change. In Keenan et al. (2021) unnormalised coprostanol concentrations and concentrations normalised to TOC were in better agreement with archaeological evidence for population change for a lake sediment core near the Maya site of Itzan than the coprostanol:(coprostanol+cholestanol)ratio. However TOC input may also vary over long time spans and this approach requires caution. The rationale for using a ratio to another stanol to take into account degradation makes sense, but for the coprostanol:(coprostanol+cholestanol) ratio to be pursued, the controls on the production and preservation of cholestanol must be further constrained. The coprostanol:(coprostanol+cholestanol) ratio is also used in contemporary sewage contamination studies, where values > 0.7 indicate sewage contamination, and values < 0.3 are representative of uncontaminated samples (Grimalt et al. 1990). Lima et al. (2019) refer to the issues of the use of this ratio in tropical areas because of diagenetic transformation of cholesterol under anaerobic conditions and cholestanol input by phytoplankton and zooplankton, and this could make the use of ratios in Central American lakes misleading. Human activities may also increase the input of nutrients to lakes resulting in phytoplankton and zooplankton blooms and it is difficult to distinguish between cholestanol from human-induced biota and cholestanol from autochthonous biota. Our data support Lima et al.’s (2019) idea, as the data point closest to Santa Elena in the Petén Itzá transect, the highest population of all lakes sampled, has a coprostanol:(coprostanol+cholestanol) ratio of 0.44 (Fig. 4e). Although this is greater than 0.3, suggesting some contamination, is not greater than the 0.7 threshold indicating sewage contamination. Some samples, with relatively low populations such as those from Lakes Macanché and Cobá (population 1150 and 1278) have coprostanol:(coprostanol+cholestanol) ratios of 0.92 and 0.99, respectively (Fig. 4a). This supports the idea that the use of normalising coprostanol to cholestanol to determine human populations can be misleading. This is further complicated by the influence of redox conditions in the sediments, and also when bottom waters are hypoxic or anoxic. In sediments, stenols are reduced to stanols, which over longer timescales are transformed into sterane, which introduces error into the ratios considered above, particularly with older samples (Wakeham, 1989). As noted, this does not necessarily preclude the use of these ratios in their use as sewage indicators or as markers of human presence of a landscape or relative population change in the past. The findings in this study of tropical lake sediments may not be universally applicable, particularly with respect to biogeochemical factors.

Conclusions

Our data from modern lakes sediments offer numerous important insights into using stanol data in order to reconstruct population history. Perhaps most importantly is that stanol concentrations are highest closest to the population centre, and this ought to be taken this into account when selecting a core site. We find much higher concentrations in the lake with the largest nearby population, Lake Petén Itzá, and concentrations clearly decrease exponentially with distance from the population centre at this lake. That all stanols decrease at a rate that approximates an exponential decline supports the idea that faecal stanols associate with particulate matter and are deposited close to shore.

Secondly, we observe that stanol concentration is correlated positively with population within 5 km from the sample location. This effect appears to be strongest for coprostanol and cholestanol, likely pointing to the fact that humans produce these stanols in large amounts, whereas stigmastanol is produced largely by herbivores, and cholesterol can also be produced in the environment. The unknown impact of reducing conditions introduces much uncertainty, and even the reduction of a small amount of ubiquitous cholesterol will strongly influence the ratios. The use of ratios can be both insightful and misleading, despite their usage in published work. The ratio coprostanol:(coprostanol+cholestanol) does not appear to correlate with human population, nor does it correlate well with proximity to population centres. The correlation of coprostanol and cholestanol in sediment raises an important question as to the source of cholestanol to lake sediments, and how this might impact its utility as a proxy for determining population remains uncertain.

Similarly, the coprostanol:(coprostanol+stigmastanol) does not correlate with human population, although it does bear some resemblance to the pattern that might be expected along the Petén Itzá transect. The coprostanol:(coprostanol+stigmastanol) appears to be useful for determining the relative contribution of humans and coprostanol producing mammals to stigmastanol producing herbivores.

The use of ratios can be complicated by the *in situ* reduction of sterols in sediments, which is difficult to quantify, or for example by waste processing, which can result in human contribution being decreased, but not the contribution of herbivore and other animals living around the lake. The use of the ratio coprostanol:(coprostanol+stigmastanol) might then be used as a guideline for assessing the dominant input of humans versus herbivores.

A way of circumventing some of the issues described above could be to normalise stanol abundance to TOC (*ESM2*), in order to take into account the effects of mineral dilution as well as the potential effects of organic matter deposition and preservation on stanol concentration. A combination of a normalising to TOC and the use of the ratio coprostanol:(coprostanol+stigmastanol) as a threshold might be the most appropriate approach in the use of stanols as a palaeo-demographic tool.

We have provided a number of new insights into the controls on stanol concentrations and ratios, and have shown that the use of stanols to reconstruct populations in the past continues to have great promise.

Acknowledgements

We thank Bjorn Sundby for his mentorship and stimulating discussions, Andrew Breckenridge for valuable feedback to this manuscript, Mark Brenner for field insights and advice, Alpacka Raft for providing a boat, Tristan Grupp for GIS troubleshooting and Thi Hao Bui for lab assistance. Funding for this project came from the Eric Mountjoy Fellowship, McGill startup funds and an NSERC Discovery Grant 2017-03902 to PMJD.

References

Battistel D, Argiriadis E, Kehrwald N, Spigariol M, Russell, JM, Barbante C (2017) Fire and human record at Lake Victoria, East Africa, during the Early Iron Age: Did humans or climate cause massive ecosystem changes? Holocene 27:997–1007

Battistel D, Piazza R, Argiriadis E, Marchiori E, Radaelli M, Barbante C (2015) GC-MS method for determining faecal sterols as biomarkers of human and pastoral animal presence in freshwater sediments. Anal Bioanal Chem 407: 8505–8514.

Bethell PH, Goad LJ, Evershed RP (1994) The study of molecular markers of human activity: The use of coprostanol in the soil as an indicator of human faecal material. J Archaeol Sci 21:619-632

Bull ID, Simpson IA, van Bergen, PF, Evershed RP (1999) Muck'n'molecules: organic geochemical methods for detecting ancient manuring. Antiquity 73:86-96

Bull ID, Elhmmali MM, Roberts DJ, & Evershed RP (2003) The application of steroidal biomarkers to track the abandonment of a Roman wastewater course at the Agora (Athens, Greece). Archaeometry 45:149–161

Constanza GR (2018) *Propuesta de plan de gestión de las aguas residuales en el área urbana de los municipios de San Benito y Flores Petén, Guatemala (Proposal* *for a wastewater management plan in the urban area of the municipalities of San Benito and Flores Petén, Guatemala). Tesis maestria en ciencias en gestion ambiental local*. Universidad San Carlos de Guatemala. Retrieved from https://www.jstage.jst.go.jp/article/amr/1/5/1_010501/_article/-char/ja/%0Ahttp://www.ghbook.ir/index.php?name=فرهنگ و رسانه های نوین&option=com_dbook&task=readonline&book_id=13650&page=73&chkhashk=ED9C9491B4&Itemid=218&lang=fa&tmpl=component%0Ahttp://dx.

D’Anjou RM, Bradley RS, Balascio NL, Finkelstein, DB (2012) Climate impacts on human settlement and agricultural activities in northern Norway revealed through sediment biogeochemistry. P Natl Acad Sci USA 109:20332–20337

Escobar J, Serna Y, Hoyos N, Velez MI, Correa-Metrio, A (2020) Why we need more paleolimnology studies in the tropics. J Paleolimnol 64: 47-53

Grimalt, JO, Fernández P, Bayona JM, Albaigés J (1990) Assessment of fecal sterols and ketones as indicators of urban sewage inputs to coastal waters. Envir Sci Tech, 24:357–363

Harrault L, Milek K, Jardé E, Jeanneau L, Derrien M, Anderson DG (2019) Faecal biomarkers can distinguish specific mammalian species in modern and past environments. PLoS One 14: e0211119

Hatcher, PG, McGillivary, PA (1979) Sewage contamination in the New York Bight. Coprostanol as an indicator. Envir Sci Tech 13:1225–1229

Hélie, JF (2009) Elemental and stable isotopic approaches for studying the organic and inorganic carbon components in natural samples. IOP C Ser Earth Env 5:012005

Hussain MA, Ford R, Hill J (2010) Determination of fecal contamination indicator sterols in an Australian water supply system. Environ Monit Assess 165:147–157

Keenan B, Imfeld A, Johnston K, Breckenridge A, Gélinas Y, Douglas PMJ (2021) Molecular evidence for human population change associated with climate events in the Maya lowlands. Quaternary Sci Rev 258:106904.

Kornilova O, Rosell-Melé A (2003) Application of microwave-assisted extraction to the analysis of biomarker climate proxies in marine sediments. Org Geochem 34: 1517–1523

LeBlanc LA, Latimer JS, Ellis JT, Quinn JG (1992) The geochemistry of coprostanol in waters and surface sediments from Narragansett Bay. Estuar Coast Shelf S 34: 439–458

Leeming R, Ball A, Ashbolt N, Nichols P (1996) Using fecal sterols from humans and animals to distinguish fecal pollution in receiving waters. Water Res 30:2893–2900

Lima MFB, Fernandes, GM, Oliveira AHB, Morais PCV, Marques EV, Santos FR, Cavalcante RM (2019) Emerging and traditional organic markers: Baseline study showing the influence of untraditional anthropogenic activities on coastal zones with multiple activities (Ceará coast, Northeast Brazil). Mar Pollut Bull 139:256–262

Lloyd CEM, Michaelides K, Chadwick DR, Dungait JAJ, Evershed RP (2012) Tracing the flow-driven vertical transport of livestock-derived organic matter through soil using biomarkers. Org Geochem 43: 56–66

Ministerio de Ambiente y Recursos Naturales (2016) Inauguran segunda fase de planta de tratamiento de aguas residuales para evitar contaminación del lago Petén Itzá. Retrieved February 10, 2021, from https://www.marn.gob.gt/noticias/actualidad/Inauguran_segunda_fase_de_planta_de_tratamiento_de_aguas_residuales_para_evitar_contaminacin_del_lago_Petn

_Itz

Murphy BT, O'Reilly SS, Monteys X, Reid BF, Szpak MT, McCaul MV, Kelleher BP (2016) The occurrence of PAHs and faecal sterols in Dublin Bay and their influence on sedimentary microbial communities. Mar Pollut Bull, 106:215–224

Pérez L, Bugja R, Lorenschat J, Brenner M, Curtis J, Hoelzmann P, Schwalb A (2011) Aquatic ecosystems of the Yucatán Peninsula (Mexico), Belize, and Guatemala. Hydrobiologia 661:407–433

Prost K, Birk JJ, Lehndorff E, Gerlach R, Amelung W (2017) Steroid biomarkers revisited - Improved source identification of faecal remains in archaeological soil material. PLoS ONE 12:1–30

Prost K, Bradel PL, Lehndorff E, Amelung W (2018) Steroid dissipation and formation in the course of farmyard manure composting. Org Geochem 118:47-57

Rodas HLS (2011) *Estudio de evaluacion de impacto ambiental: Diseño de la fase II de la planta de tratamiento de aguas residuales del area central ubicada en San Benito, Departmento de Peten. Ministerio de Ambiente y Recursos Naturales (MARN).*

Sarkar S, Wilkes H, Prasad S, Brauer A, Riedel N, Stebich M, Sachse D (2014) Spatial heterogeneity in lipid biomarker distributions in the catchment and sediments of a crater lake in central India. Org Geochem 66:125–136

Sear DA, Allen MS., Hassall JD, Maloney AE, Langdon PG, Morrison AE, Pearson E (2020) Human settlement of East Polynesia earlier, incremental, and coincident with prolonged South Pacific drought. P Natl Acad Sci USA 117:13846

Shillito LM, Whelton HL, Blong JC, Jenkins DL, Connolly TJ, Bull ID (2020). Pre-Clovis occupation of the Americas identified by human fecal biomarkers in coprolites from Paisley Caves, Oregon. Sci Adv 6:1–9

Sistiaga A, Berna F, Laursen R, Goldberg, P (2014) Steroidal biomarker analysis of a 14 , 000 years old putative human coprolite from Paisley Cave , Oregon. J Archaeol Sci 41:813–817

Speranza ED, Colombo M, Heguilor S, Tatone LM, Colombo JC (2020) Alterations in the sterol signature of detritivorous fish along pollution gradients in the Río de la Plata basin (Argentina): From plant to sewage-based diet. Environ Res 184:109351

Teshima S, Kanazawa A (1978) Conversion of cholesterol to coprostanol and cholestanol in the estuary sediment. Mem Fac Fish, Kagoshima Univ 27:1–47.

Thienemann M, Masi A, Kusch S, Sadori L, John S, Francke A, Rethemeyer, J (2017) Organic geochemical and palynological evidence for Holocene natural and anthropogenic environmental change at Lake Dojran (Macedonia/Greece). Holocene 27:1103–1114

Vachula RS, Huang, Y, Longo WM, Dee SG, Daniels WC, Russell JM (2019) Evidence of Ice Age humans in eastern Beringia suggests early migration to North America. Quaternary Sci Rev 205: 35–44

Venkatesan MI, Ruth, E, Kaplan IR (1986) Coprostanols in Antarctic marine sediments: A biomarker for marine mammals and not human pollution. Mar Pollut Bull 17:554–557

Venkatesan MI, Santiago, CA (1989) Sterols in ocean sediments: novel tracers to examine habitats of cetaceans, pinnipeds, penguins and humans. Mar Biol 102: 431–437

Wakeham SG (1989) Reduction of stenols to stanols in particulate matter oxic-anoxic boundaries in sea water. Nature *342*:787–790

White AJ, Stevens LR, Lorenzi V, Munoz SE, Lipo CP, Schroeder S (2018) An evaluation of fecal stanols as indicators of population change at Cahokia, Illinois. J Archaeol Sci 93:129–134

Zhang Y, Zhang D, Yang Y, Wu X (2020) Pollen and lipid analysis of coprolites from Yuhuicun and Houtieying, China: Implications for human habitats and diets. J Archaeol Sci Rep 29:102135

Figure captions

Figure 1. a) Map showing location of study area in Central America; b) map of sampling sites on the Yucatán Peninsula with the Petén department in northern Guatemala; and c) locations of samples collected along transect in Lake Petén Itzá in detail.

Figure 2. Scatter plot of total organic carbon and stanols (a) coprostanol; b) stigmastanol; c) cholestanol; and d) cholesterol (Ng/g dry sediment) for Petén Itzá (red) and all other lakes (blue).

Figure 3. Relationships between concentrations of coprostanol+epicoprostanol and other stanols. Red data points and best-fit line for Petén Itzá, blue for all lakes excluding Petén Itzá, and the grey best-fit line is for all lakes. a) coprostanol and cholestanol; b) coprostanol and stigmastanol; c) coprostanol and cholesterol; d) cholestanol and cholesterol; e) cholesterol and stigmastanol; and f) cholestanol and stigmastanol. For stigmastanol regressions the outlier has not been included in the regression.

Figure 4. a) Relationship between coprostanol concentrations and population for all lakes excluding Lake Petén Itzá; b) Relationship between coprostanol concentrations and population for all lakes excluding Lake Petén Itzá (points are from sampling locations closest to the shore); b) for stigmastanol; c) for cholestanol; d) for cholesterol; e) for coprostanol:(coprostanol+cholestanol); and f) for coprostanol:(coprostanol+stigmastanol). The green circles are data points from lakes with high salinity.

Figure 5. Changes in absolute stanol concentrations or stanol ratios along a 6.5 km transect from Santa Elena, in Lake Petén Itzá for: a) coprostanol; b) stigmastanol; c) cholestanol; d) cholesterol; e) coprostanol:(coprostanol+cholestanol); and f) coprostanol:(coprostanol+stigmastanol). Stanol concentrations normalised to OC are included in *ESM2*. Note that at 1 km the sample with low stanol concentrations is on the other side of the island of Flores.

Figures

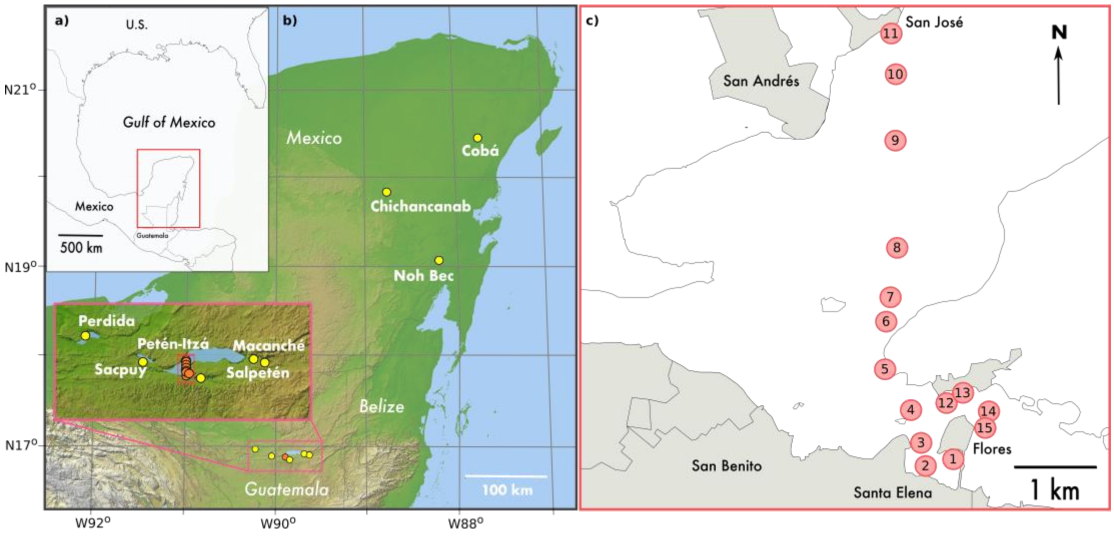

Figure 1.

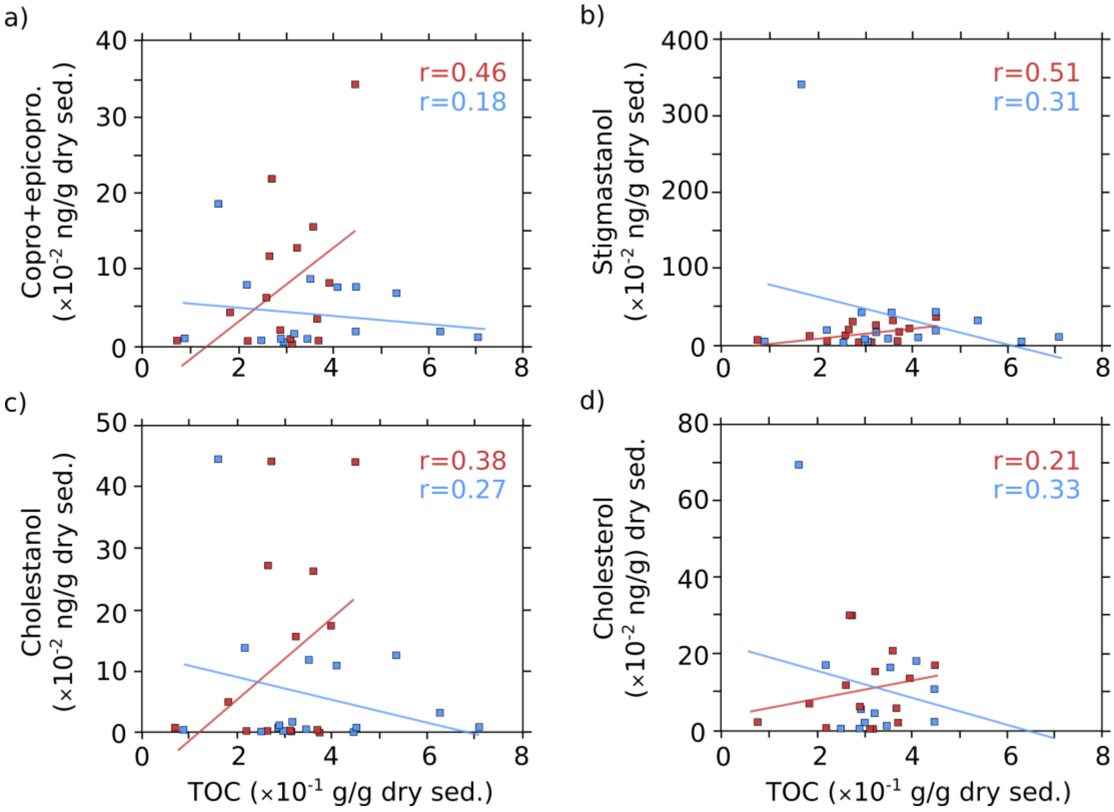

Figure 2.

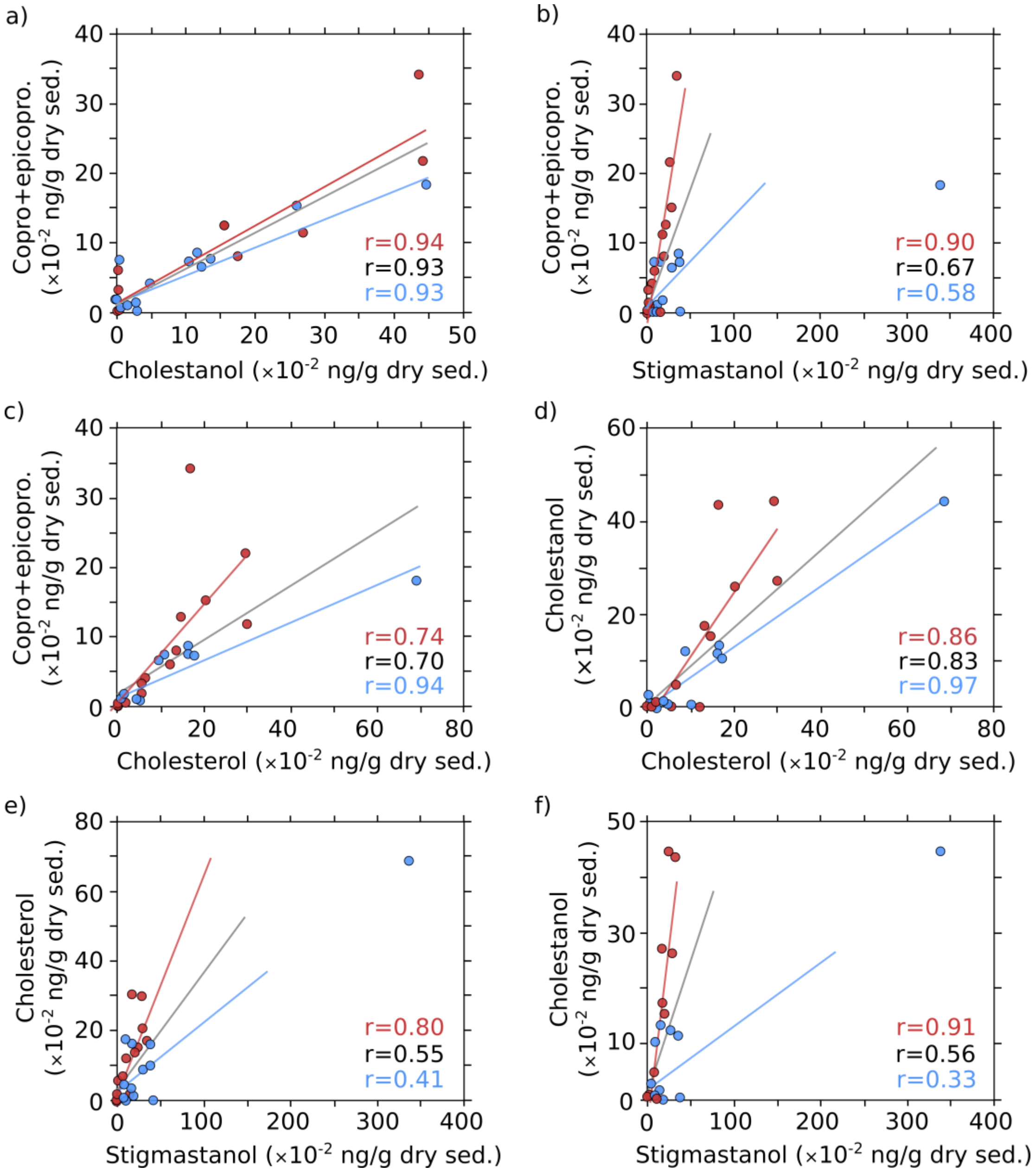
a)
Copro+epicopro. (×10^{-2} ng/g dry sed.)
Cholestanol (×10^{-2} ng/g dry sed.)
0
10
20
30
40
50
r=0.94
r=0.93
r=0.93
b)
Copro+epicopro. (×10^{-2} ng/g dry sed.)
Stigmastanol (×10^{-2} ng/g dry sed.)
0
100
200
300
400
r=0.90
r=0.67
r=0.58
c)
Copro+epicopro. (×10^{-2} ng/g dry sed.)
Cholesterol (×10^{-2} ng/g dry sed.)
0
20
40
60
80
r=0.74
r=0.70
r=0.94
d)
Cholestanol (×10^{-2} ng/g dry sed.)
Cholesterol (×10^{-2} ng/g dry sed.)
r=0.86
r=0.83
r=0.97
e)
Cholesterol (×10^{-2} ng/g dry sed.)
Stigmastanol (×10^{-2} ng/g dry sed.)
r=0.80
r=0.55
r=0.41
f)
Cholestanol (×10^{-2} ng/g dry sed.)
Stigmastanol (×10^{-2} ng/g dry sed.)
r=0.91
r=0.56
r=0.33

Figure 3.

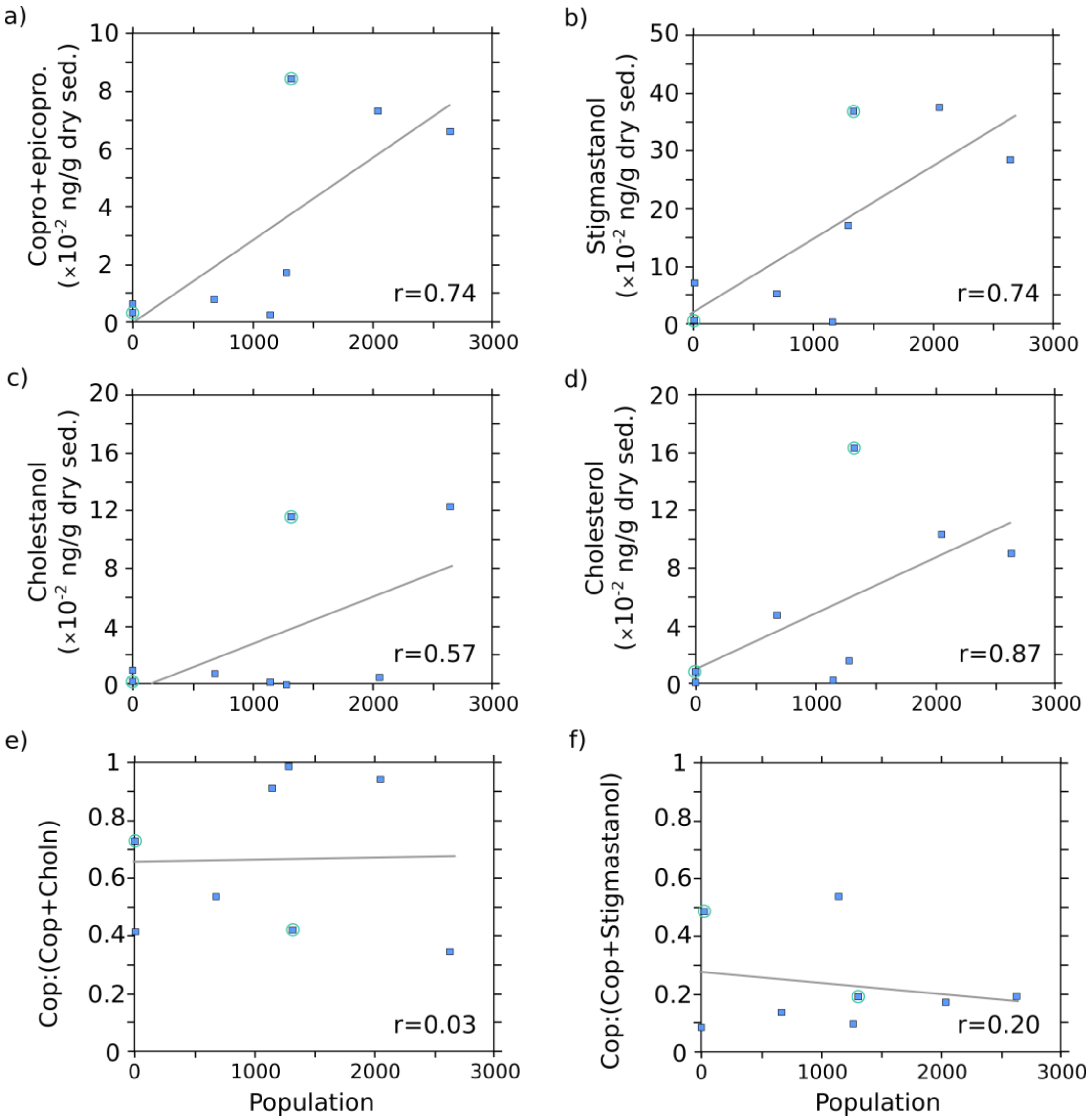
a)
Copro+epicopro.
(×10^{-2} ng/g dry sed.)
0
2
4
6
8
10
0
1000
2000
3000
r=0.74
b)
Stigmastanol
(×10^{-2} ng/g dry sed.)
0
10
20
30
40
50
0
1000
2000
3000
r=0.74
c)
Cholestanol
(×10^{-2} ng/g dry sed.)
0
4
8
12
16
20
0
1000
2000
3000
r=0.57
d)
Cholesterol
(×10^{-2} ng/g dry sed.)
0
4
8
12
16
20
0
1000
2000
3000
r=0.87
e)
Cop:(Cop+Choln)
0
0.2
0.4
0.6
0.8
1
0
1000
2000
3000
Population
r=0.03
f)
Cop:(Cop+Stigmastanol)
0
0.2
0.4
0.6
0.8
1
0
1000
2000
3000
Population
r=0.20

Figure 4.

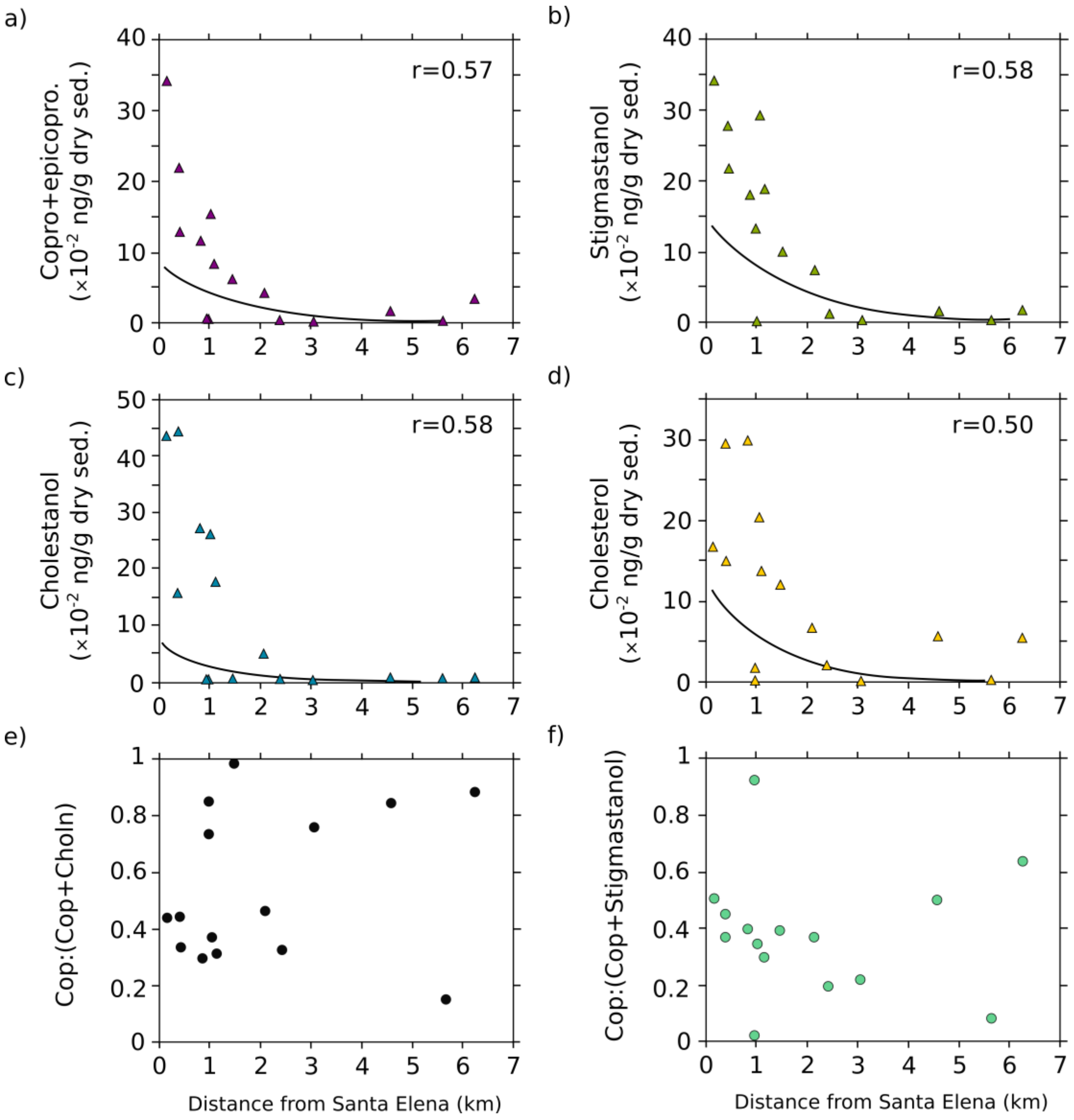

Figure 5.

Tables

Table I. Lake sediment sample locations, water depths, %Corg and stanol and sterol concentrations. Petén Itzá samples are ordered by the location along the transect (distance from Santa Elena). Samples with YSI measurements are indicated by an asterisk (*).