

scientific data



OPEN

DATA DESCRIPTOR

Demographic, behavioral, and ecological data from a long-term field study of wild baboons in Amboseli, Kenya

Chelsea A. Southworth^{1,20}✉, Jack C. Winans^{2,3,20}✉, Jacob B. Gordon⁴, Niki H. Learn⁵, William A. Wilber¹, Catherine Andreadis¹, Gretchen Andreasen¹, Mimi Arandjelovic^{6,7}, C. Ryan Campbell⁸, Mary N. Chege^{1,9,10}, Maria J. A. Creighton⁴, Carmen M. Cromer⁴, Reena Debray⁶, Carly C. Dickson¹, Pamela Ferretti¹¹, Elizabeth M. George¹², Laurence R. Gesquiere⁴, Shuyu He^{4,6}, Leif Hey¹³, Emily E. Jefferson¹⁴, Ipek G. Kulahci¹, Brian A. Lerch¹⁵, Lee Nonnamaker¹⁶, Iker Rivas-González⁶, Beniamino Tulliozi⁴, Shasta E. Webb^{1,17}, Susan C. Alberts^{4,13,20}, Elizabeth A. Archie^{1,20} & Jenny Tung^{4,6,13,18,19,20}

Long-term data sets on individually recognized animals and their environments are critical to understanding animal behavior, evolution, and ecology. However, they are resource- and time-intensive and seldom made publicly available. The Amboseli Baboon Research Project (ABRP) is one of the longest-running studies of a wild mammal population in the world and has collected extensive data on the baboon population of the Amboseli ecosystem in Kenya since 1971. Here, we describe four ABRP data sets newly available to the evolutionary biology, behavioral ecology, and primatology communities: (1) the sizes and demographic compositions of 21 social groups from 1971–2023; (2) the activity budgets of adult females and immatures from 1984–2023; (3) behavioral data on diet for adult females and immatures from 1984–2023; and (4) weather data, including precipitation from 1976–2023 and temperature from 1976–2022. Data are aggregated annually and monthly to enable cross-data set analyses. These data offer a rare longitudinal perspective on behavioral and ecological change in a wild mammal population.

¹Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, USA. ²Department for the Ecology of Animal Societies, Max Planck Institute of Animal Behavior, Konstanz, Germany. ³Department of Biology, University of Konstanz, Konstanz, Germany. ⁴Department of Biology, Duke University, Durham, NC, USA. ⁵Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA. ⁶Department of Primate Behavior and Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. ⁷German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany. ⁸Biostatistics & Computational Biology Branch, National Institute of Environmental Health Sciences, Durham, NC, USA. ⁹One Health Centre, Kenya Institute of Primate Research, Nairobi, Kenya. ¹⁰Department of Biochemistry, University of Nairobi, Nairobi, Kenya. ¹¹Section of Genetic Medicine, Department of Medicine, University of Chicago, Chicago, IL, USA. ¹²Department of Biology, Gannon University, Erie, PA, USA. ¹³Department of Evolutionary Anthropology, Duke University, Durham, NC, USA. ¹⁴Interdepartmental Doctoral Program in Anthropological Sciences, Stony Brook University, Stony Brook, NY, USA. ¹⁵Center for Population Biology, University of California, Davis, CA, USA. ¹⁶Department of Biology, University of Florida, Gainesville, FL, USA. ¹⁷Department of Anthropology and Archaeology, University of Calgary, Calgary, AB, Canada. ¹⁸Canadian Institute of Advanced Research, Toronto, ON, Canada. ¹⁹Faculty of Life Sciences, Leipzig University, Leipzig, Germany. ²⁰These authors contributed equally: Chelsea A. Southworth, Jack C. Winans, Susan C. Alberts, Elizabeth A. Archie, Jenny Tung. ✉e-mail: csouthwo@nd.edu; jwinans@ab.mpg.de

Background

Longitudinal data on wild animal populations are essential to understanding the relationships between animal behavior, ecology, and fitness. Amassing long-term data sets that span the lives of many animals allows researchers to investigate these relationships with considerable power^{1,2}. However, life course data of this scale are rare for long-lived species³, even though cross-sectional data often fail to capture the full range of social and environmental conditions that individuals experience⁴. To spur methods development, facilitate comparative analyses, and understand change over time, long-term data sets on relevant behavioral, demographic, and environmental variables are required, and, where feasible, should be made publicly available.

The Amboseli Baboon Research Project (ABRP) is one of the longest-running field studies of a wild mammal population in the world. Founded in 1971 by Jeanne and Stuart Altmann, the ABRP has observed the baboon population living in the Amboseli basin for more than 50 years^{4–6}. Baboons are highly social, ecologically flexible primates that are widely distributed across Africa and the Arabian peninsula⁷. In Amboseli, baboons live in stable social groups of varying sizes (~20 to over 120 individuals), consisting of multiple adult females and males as well as immatures of both sexes⁶. Females are philopatric, while males typically disperse to other social groups at sexual maturity and can move between groups multiple times in their lives^{8–12}. The baboons in this ecosystem, like many baboon populations across Africa, experience highly variable environments, including pronounced seasonal variation and high yearly variance in temperature and rainfall conditions, even within season^{13,14}. They cope with this variability through behavioral flexibility—adjusting activity budgets and capitalizing on seasonal food availability as highly selective ‘eclectic omnivores’^{7,13,15–17}. These ecological and dietary pressures in turn shape many aspects of the baboons’ lives, including their maturation and reproductive schedules, body condition, and movement patterns^{13,15,17–21}.

The ABRP conducts intensive data collection on multiple social groups with overlapping home ranges, alongside opportunistic monitoring of other groups that range in the study area^{4,5}. Over the course of the project, more than 2,000 individually-recognized animals have been monitored⁶. Here, we describe and make newly available four of the resulting long-term data sets: 1) group demography data for 21 social groups that were followed for different periods over 52 years, 2) activity budgets of adult females and immatures of both sexes over 39 years, 3) diet composition at the social group and study population levels over 39 years, and 4) rainfall and temperature measurements over 47 years²². Each data set is aggregated over the same timescales (months and years), facilitating cross-data set analyses. We provide visual summaries of patterns across time and social groups for each dataset, as well as cleaned, aggregated, validated, and processed data for other researchers to incorporate in their own analyses. We hope that these data will facilitate comparative analyses of animal behavior and behavioral and ecological change over time and may be useful as a teaching tool. In addition, this repository will be useful for the many researchers working with ABRP data or other east African mammals as an open-access reference for demographic, behavioral, and ecological patterns in this ecosystem, from 1971 to 2023.

Methods

Study site and subjects. Since 1971, the ABRP has studied a naturally admixed population of yellow (*Papio cynocephalus*) and anubis baboons (*P. anubis*; most population members are of predominantly *P. cynocephalus* ancestry^{23,24}) located in the Amboseli ecosystem of southern Kenya (2°40’S, 37°15’E, 1100 m altitude)⁴. Amboseli is a semi-arid, short grass savannah ecosystem that experiences considerable seasonal and inter-annual variation in precipitation and temperature^{13,14} (Fig. 1). Generally, four seasons of varying lengths characterize intra-annual rainfall patterns in Amboseli¹⁴. The hydrological year begins with a short wet season lasting from November to January, followed by a short dry season in February, a long wet season from March to May, and a long dry season from May to October (Fig. 1a). In all data sets provided here, the hydrological year therefore runs from November 1 to October 31 (e.g., hydrological year 2023 encompasses November 1, 2022 to October 31, 2023). Daily temperatures peak in February and March during the short dry and early long wet season and fall to their lowest in the long dry season from June to August (Fig. 1b). Rain almost never falls in the long dry season, but Amboseli experiences substantial inter-annual variability in rainfall during the rest of the hydrological year (Fig. 1c). Maximum and minimum temperatures vary inter-annually, but are more consistent compared to rainfall (Fig. 1d).

Over the course of the long-term study, the ABRP has studied between one and six social groups at a time, all descendants through fission or fusion of two original study groups on which observation began in 1971 and 1980, respectively⁴ (Fig. 2). Throughout the study period, data were consistently collected on demography, behavior, and reproductive patterns; collection of biological samples for genetic and hormone analysis began in the 1980s and 1990s, respectively⁴. We distinguish between social groups that were regularly subject to such intensive data collection (‘study groups’) and other social groups that were present in the ecosystem but were monitored on an irregular or infrequent basis (‘non-study groups’). The number of study groups has varied over time due to the number of project observers, changes in the size of the population, and permanent fissions or fusions (Fig. 2). To maintain high levels of data collection effort across study groups, some study groups were dropped from intensive data collection over time (Fig. 2). Detailed descriptions of the data collection protocols maintained by ABRP can be found online at the ABRP’s website²⁵.

Demographic and behavioral data collection. The group demographic data were compiled from census records collected between 1971 and 2023, while the activity budget and diet data came from focal animal samples conducted between 1984 and 2023. Experienced field researchers performed group censuses and focal animal samples in the context of half-day visits to each group, which were conducted in either the morning or afternoon, 6 days per week. At this schedule, the typical frequency with which each group was observed ranged from 2–4 days per week to near-daily, depending on the number of groups being monitored at a time and the number of project observers. Over the history of the project, the mean number of observers monitoring each group in a month was 2.75 (range: 1 – 5). Information on observers was only available when behavioral data were collected.

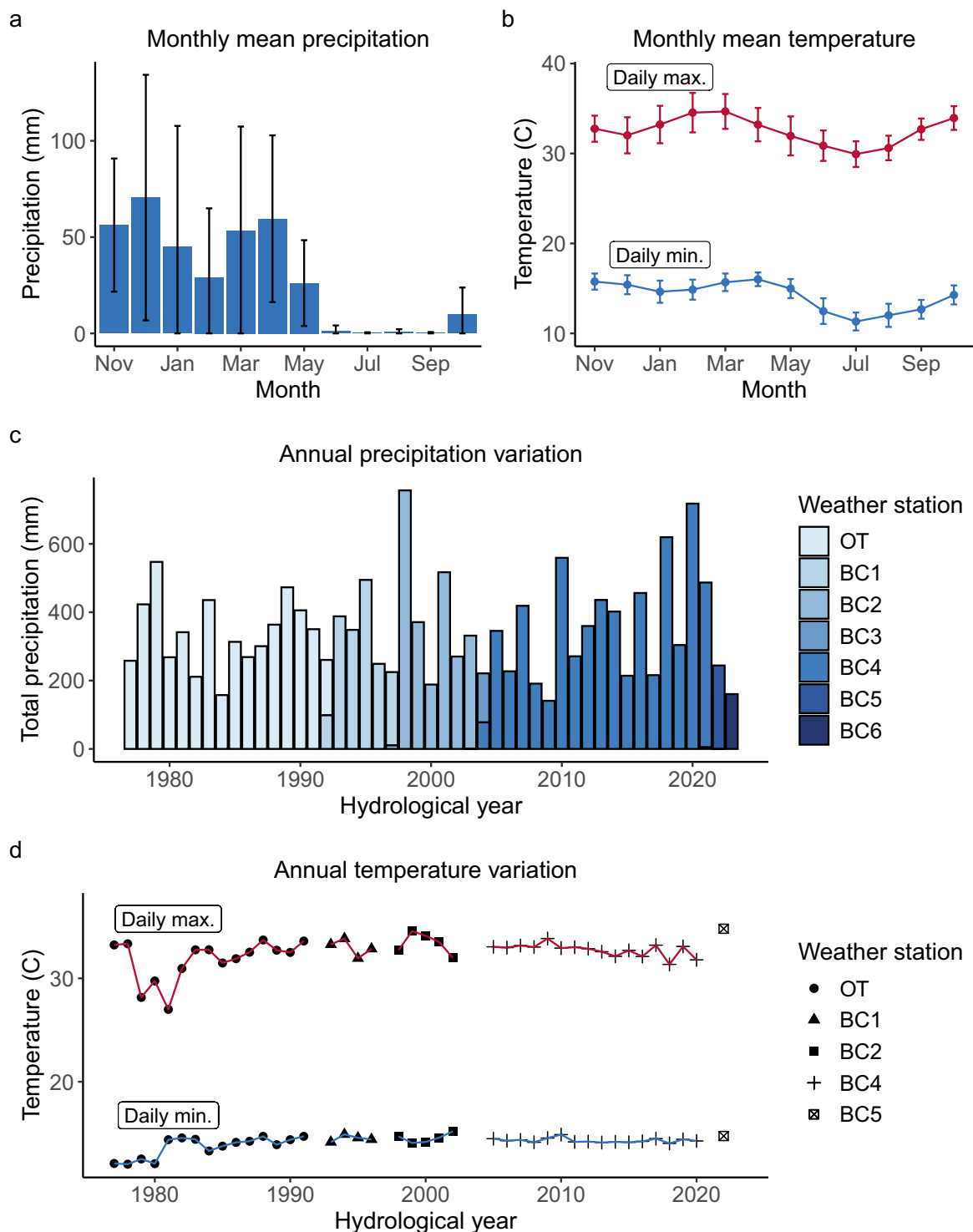


Fig. 1 Monthly and annual patterns of precipitation and temperature in Amboseli, Kenya. Temperatures were taken from analog min-max thermometers and precipitation measures were taken from a standard rain gauge. The location of the thermometers and rain gauges changed slightly over the years but, with one exception, were within 150 m of each other. **(a,b)** depict monthly patterns over the hydrological year, which runs from November to October. In **(a)**, bar heights represent mean precipitation volume in each month and error bars represent standard deviations within months between hydrological years 1977 and 2023. In **(b)**, points represent mean daily maximum (red) and minimum (blue) temperatures and error bars represent standard deviations within months between hydrological years 1977 and 2022. **(c)** depicts total annual precipitation from hydrological years 1977 to 2023. **(d)** depicts annual average minimum and maximum temperature from hydrological years 1977 to 2022. In **(c)**, the height of the bars represents the total amount of rainfall in each hydrological year. In **(d)**, the black points on the red and blue lines depict the mean maximum and minimum daily temperature, respectively, in each hydrological year. Annual mean temperature values were only included in **(d)** if all data for a given hydrological year were collected by the same weather station.

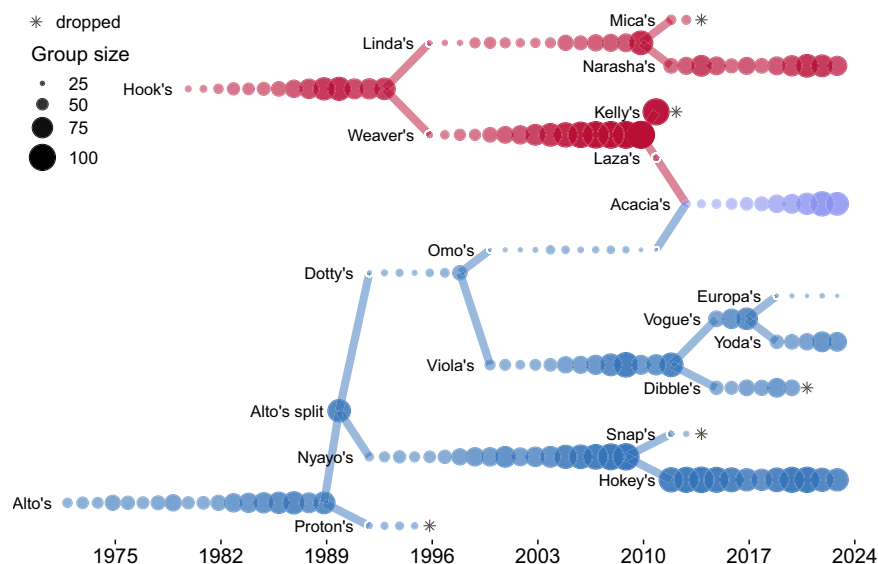


Fig. 2 Total size of social groups subject to intensive monitoring by the Amboseli Baboon Research Project over time. Each group is represented by one data point for each hydrological year that it was intensively monitored and included in the study. Point size varies by group size, and lines connecting groups represent fission events (when groups permanently split) and fusion events (when groups permanently merged). Alto's group, one of the two original study groups, and its daughter groups are represented in blue, and Hook's group, the other original study group, and its daughter groups are represented in red. Acacia's group was formed by the fusion of one daughter group from each lineage and is represented in purple. Asterisks indicate instances in which a group was dropped from intensive data collection.

If only a group census was conducted, e.g. due to time constraints or disruption, no observer was recorded and the observer count column will contain 'NA.'

Census data were collected at the start of each group visit, including information on the number and identity of all individuals present in the group and on any demographic events, including births, known or suspected deaths, and known or suspected dispersal events. From these data, we calculated, for each social group in each month and hydrological year: i) the total number of unique individuals that were resident in the group, including ii) the total number of unique adult females, iii) the total number of unique adult males, and iv) the total number of unique non-adult individuals. We used the concept of 'residency' to distinguish stable group membership from unstable 'floating' between groups, which was generally seen only in adult males but sometimes happened with other age-sex classes during group fissions. In principle, an individual was considered 'resident' in a given social group over a given period of time if they were present in the group more frequently than they were absent from it. In practice, residency was calculated algorithmically for each individual on each day via day-by-day assessments of their census observations (for full details, see²⁶). For the yearly demographic data set²², any female who had reached menarche before or during a given hydrological year was considered an adult in that year and any male who had attained a dominance rank above at least one other adult male before or during a given hydrological year was considered an adult in that year^{27,28}. Analogously, for the monthly demographic data set²², individuals who reached adulthood before or during a given month were counted as adults for that month.

These demographic data²² were collected from 21 unique social groups, including 185 complete group-years and 31 partial group-years, and 2,372 group-months. Partial group-years occurred when groups fissioned or fused, or when logistical obstacles interfered with intensive data collection on a group in a given year. When a group underwent a permanent fission or when two groups fused, the products of the fission/fusion were considered new social groups, i.e., no group retained an original group's 'identity.'

During group visits, focal animal samples were conducted on adult females and on juveniles of both sexes in a pre-determined, randomized order²⁹. Adult males were not regular subjects of focal animal sampling effort as part of the long-term research project. Focal animal samples and the resulting activity budgets and diet data were typically collected between 0700 and 1100 or between 1400 and 1800. Samples lasted 10 minutes and on each minute during the sample, the focal animal's activity (feeding, walking, resting, or socializing) was recorded (Fig. 3). Beginning in September 1985, observers began additionally recording the type of food focal individuals were eating when the activity was 'feeding' (Fig. 4). Observers attempted to assign all diet items to one of 73 types, each of which possesses a unique three-letter identification code. For some food types (e.g., diet items from *Vachellia* trees), the species name and plant part were encoded in this assignment, but for others (e.g., diet items from some grasses), species identification was not consistently possible. Therefore, we also categorized diet items into eight categories based on phylogenetic and/or structural properties that together make up 93.5% of total time spent feeding during years when food types were identified (see Fig. 4). We consider these eight categories to vary meaningfully with respect to nutritional value, processing demands, spatio-temporal

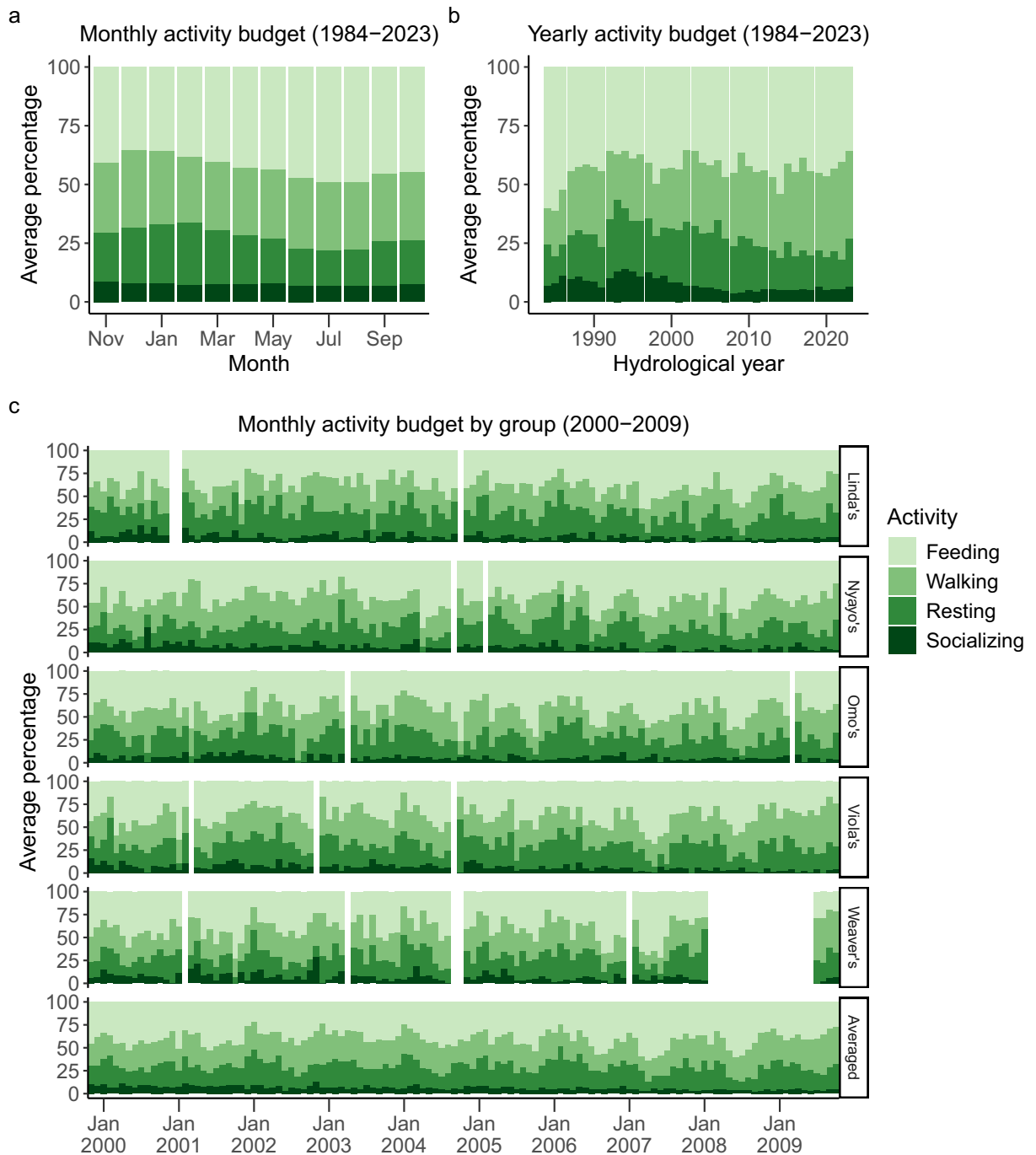


Fig. 3 Variation in the percentage of focal sample time that adult females spent feeding, walking, resting, and socializing (a) in each month, averaged across study groups and years, (b) within each year from 1984–2023, averaged across study groups, and (c) within five social groups, averaged within months from November 1999 to October 2009 (i.e., hydrological years 2000–2009) and averaged within months across the five social groups over the same time period (‘Averaged’ panel). ‘Socializing’ included time that the focal individual spent grooming, being groomed, and involved in other social interactions with conspecifics. In (c), white spaces indicate group-months in which we believe focal animal sampling was too sparse to produce a statistically representative sample of the activity budgets ($n < 10$ focal animal samples).

dispersion, and other characteristics of biological significance to foraging baboons¹⁷. Of these categories, we consider grass corms, grass blade bases and seed heads, and *Vachellia* seeds to be low-energy foods, and we consider grass blades, fruits, flowers, tree gum, and invertebrates to be high-energy foods^{14,17}. The diet items that constituted the remaining 6.5% of feeding time, including unidentifiable diet items, were classified into a ninth category called ‘Other’. Because observers did not record any information about food types during focal animal

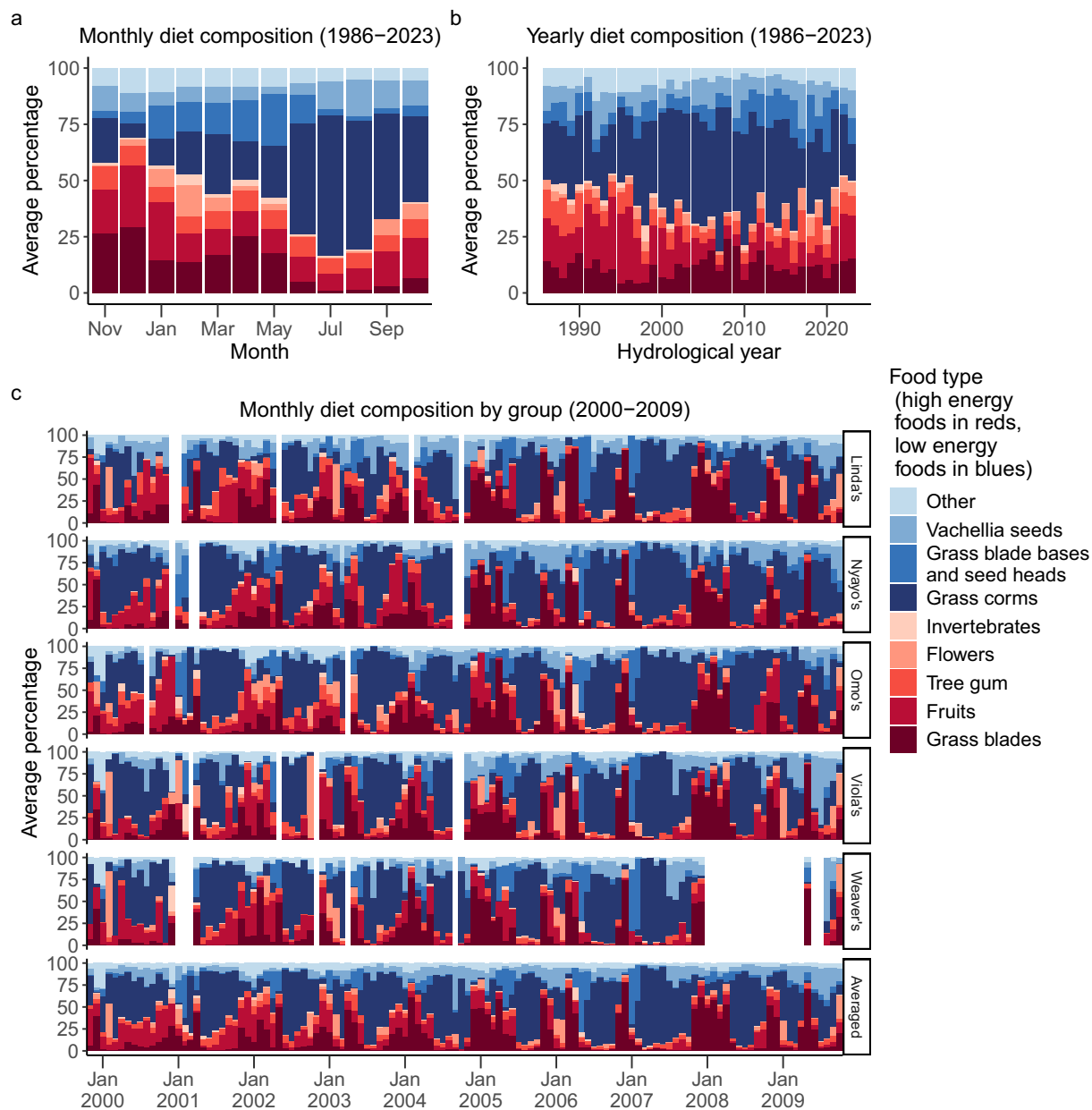


Fig. 4 Variation in the percentage of total feeding time spent on each major food type (a) in each month, averaged across years and across study groups, (c) within each year from 1986–2023, averaged across study groups, and (c) within five study groups, averaged within months from November 1999 to October 2009 (i.e., hydrological years 2000–2009) and averaged within months across the five social groups over the same time period ('Averaged' panel). Presumed high-energy food types are represented in shades of red and low-energy food types are represented in shades of blue. In (c), white spaces indicate group-months in which we believe that focal animal sampling was too sparse to produce a statistically representative sample of diet composition ($n < 75$ minutes of feeding observed).

sampling from January 1984 through August 1985, all feeding observations from this period were labelled as 'not recorded' food types.

These activity budget and diet data²² were collected from 21 unique social groups across 161–180 group-years and 1,682–1,905 group-months. Because we did not collect activity budget and diet data prior to 1984, rows for those years contain 'NA.' After 1984, 'NA' values could result from several reasons, including: observations were attempted, but no focal samples contained in-sight observations; focal samples were not collected due to in-progress group fissions or fusions; or other logistical obstacles prevented detailed focal sampling.

Weather monitoring. The fourth data set published here (weather data)²² resulted from measurements of daily minimum and maximum temperature and daily rainfall from analog min-max thermometers and rain gauges kept in our field camp. In all years, the thermometer and rain gauge were within a few meters of each other

in an area that we term our ‘weather station’. However, the location of the weather station changed over the years, once because we moved our field camp approximately 7.1 km (in 1992), and after that because we moved the weather station to various locations within our field camp as the layout of camp changed (distances between subsequent locations were generally less than 150 m). We include a unique identifier to indicate the weather station where the data were collected and the location (latitude/longitude) of each station. These data were recorded from 1976 to 2022 for temperature and 1976 to 2023 for rainfall. From these daily measurements, we calculated for each month and year: (i) the average minimum daily temperature, (ii) the average maximum daily temperature, and (iii) the total volume of rainfall.

Ethics statement. This research was approved by the IACUC at Duke University, University of Notre Dame, and the Ethics Council of the Max Planck Society and adhered to all laws and guidelines of Kenya.

Data Records

The data²² are accessible on Dryad (<https://doi.org/10.5061/dryad.wwpzgmzsf>). Descriptions for each column in each data record are included in `table_descriptions.xlsx`.

Social group history. Aggregated metadata for all ABRP study groups observed for any length of time. File `social_group_history.csv` (21 social groups) provides information on time periods of observation and social group fissions and fusions for each study group. These metadata are essential for interpreting the group-level data summaries below.

Weather station locations. Latitudes and longitudes for each of the 7 weather stations from which weather data were collected between 1976 and 2023 in file `weather_station_locations.csv`. Exact coordinates were not recorded for stations BC2 and BC3.

Monthly data for each study group. Study group composition, activity budget, and diet data are aggregated by month and social group in file `monthly_grp_comp_activity_diet.csv` (2,372 total group-months). Columns A-C include the social group name, year, and month of data presented in a given row.

Study group composition data (2,372 group-months of data) are included in columns D-I. These data present total group size, numbers of resident adult females, adult males, and immatures, whether the group was undergoing a fission or fusion, and any group-level notes on demography for a social group during a given month from 1971 to 2023.

Adult female activity budget data (1,901 group-months of data) are included in columns J-U. These data present adult female activity budgets (i.e., activity budgets aggregated across all adult females in the group) for each study group during each month and were collected from focal animal samples on adult females between 1984 and 2023. Data were aggregated by pooling all data points collected at one-minute intervals during focal samples, across all adult females in the group, resulting in a group-level estimate of the proportion of total sampling time that the pooled set of adult females spent in each activity (or on each food type).

Immature animal activity budget data (1,682 group-months of data) are included in columns V-AI. These data present immature activity budgets (i.e., activity budgets aggregated across all immature individuals in the group, as described above for adult females) for each study group during each month and were collected from focal animal samples on immatures between 1984 and 2023.

Diet data (1,905 group-months of data) are included in columns AJ-BD. These data present the proportion of feeding time that adult females and immatures spent on different food categories for each study group during each month and were collected from focal animal samples on adult females and immatures between 1984 and 2023. Prior to September 1985, observers did not distinguish between different food categories when collecting focal animal samples.

Observer data (2,176 group-months of data) are included in column BE. These data sum the total number of unique observers who collected behavioral data from each group during each month. These observers also conducted group censuses.

Monthly data aggregated across all groups. Activity budget and diet data are aggregated across all study groups and weather data are aggregated by month in the file `monthly_activity_diet_weather.csv` (569 total months). Here, data were aggregated by pooling all data points collected at one-minute intervals during focal samples, across all adult females in all study groups, resulting in a population-level estimate of the proportion of total sampling time that the pooled set of adult females in all study groups spent in each activity (or on each food type). Columns A-B include the year and month of data summarized in a given row. In three of the months shown in the table, data were collected from two weather stations, and each station has a separate row: June 1997, when BC1 was transitioning to BC2; February 2004, when BC3 was transitioning to BC4; and June 2021, when BC4 was transitioning to BC5.

Adult female activity budget data (475 months of data) are included in columns C-N. These data summarize adult female activity budgets during a given month, aggregated across all study groups, and were collected from focal animal samples on adult females between 1984 and 2023.

Immature animal activity budget data (376 months of data) are included in columns O-AB. These data summarize immature activity budgets during a given month, aggregated across all study groups, and were collected from focal animal samples on immatures between 1984 and 2023.

Diet data (478 months of data) are included in columns AC-AW. These data present the proportion of feeding time that adult females and immatures spent on different food categories during a given month, aggregated across all groups, and were collected from focal animal samples on adult females and immatures between 1984

and 2023. Prior to September 1985, observers did not distinguish between different food types when collecting focal animal samples.

Weather data (569 months of data) are included in columns AX-BA. These data include the average daily temperature minimum, temperature maximum, and total rainfall for each month and were collected from a series of different weather stations between 1976 and 2023.

Yearly data for each study group. These data parallel the monthly data for each study group described above but are aggregated at the level of the hydrological year. In this case, data were aggregated by pooling all data points collected in each study group during the entire hydrological year, resulting in an annual group-level estimate of the proportion of total sampling time that the pooled set of adult females spent in each activity (or on each food type). Social group composition, activity budget, and diet data are aggregated by hydrological year and study group in file `hydroyear_grp_comp_activity_diet.csv` (216 total group-years). Columns A-B include the study group name and hydrological year of data presented in a given row.

Study group composition data (216 group-years of data) are included in columns C-H. These data present total group size, numbers of resident adult females, adult males, and immatures, whether the group was undergoing a fission or fusion, and group-level demography notes for a social group during a given hydrological year from 1971 to 2023.

Adult female activity budget data (180 group-years of data) are included in columns I-T. These data present adult female activity budgets for a study group during a given hydrological year and were collected from focal animal samples on adult females between 1984 and 2023.

Immature animal activity budget data (161 group-years of data) are included in columns U-AH. These data present immature activity budgets for a study group during a given hydrological year and were collected from focal animal samples on immatures between 1984 and 2023.

Diet data (180 group-years of data) are included in columns AI-BC. These data summarize adult female and immature diets for a study group during a given hydrological year and were collected from focal animal samples on adult females and immatures between 1984 and 2023. Prior to September 1985, observers did not distinguish between different food categories when collecting focal animal samples. All diet data for hydrological year 1985 were categorized as ‘Unclassified’ in this sheet as distinct food categories were only collected for in September and October.

Yearly weather data. Weather data are aggregated by hydrological year in file `hydroyearly_weather_data.csv` (47 years and 54 rows of data). These data include average daily temperature minimum and average daily temperature maximum assessed over each hydrological year, and total rainfall for each hydrological year. They were collected from a series of different weather stations between 1976 and 2023.

Detailed diet data. Detailed diet data that specify all 73 food types we collected (i.e., the unaggregated form of the 10 food categories presented in Fig. 4) are aggregated by month and social group in file `detailed_diet.csv` (1,905 group-months of data). Each row is associated with a unique social group, year, month, and one of the 73 different food types. There is only a row for a given food type if it was observed being eaten by a member of that unique social group during a focal sample collected in that year and month. These data present the number of observations of adult females and immatures feeding on different food types during a given month in a given social group and were collected during focal animal sampling between 1984 and 2023. Prior to September 1985, observers did not distinguish between different food types when collecting focal animal samples.

Technical Validation

We strive to validate the data we collect by implementing various precautions, safety checks, and database management practices. Detailed explanations of these procedures are available online^{25,26}. In the field, skilled observers recorded focal animal sample data as handwritten records between January 1984 and 1999, and electronically beginning in August 1999 (with Psion Workabout units through July 2015 and thereafter with the Prim8 Android app on Samsung tablets³⁰). Focal data were sent to database managers in the United States monthly between 1984 and approximately 2005, and on a weekly basis from 2005 until 2015; focal data recorded since 2015 have been sent daily. Handwritten data, including census and weather data, were proofread in the field every Saturday by multiple field team members, who searched for potential inconsistencies or incomplete data. At the end of each month, all handwritten data were proofread again, and (since 2014) scanned and emailed to database managers in the United States; before 2014, handwritten data were mailed to the United States each month. Additionally, on the last four observation days of each month, the field observers collected a reduced number of focal samples in order to collect other types of data, including wounds and pathologies and signs of physical maturation in males. They also used this time to ensure that different observers were reliably identifying the same individuals, which is particularly important for maturing infants and juveniles who can experience rapid changes in coat color and size.

Upon the arrival of the data in the United States, database managers sorted and proofread the data again, discussing any potential errors with the field team. Data were ultimately uploaded to our PostgreSQL database, Babase^{31,32}. Babase has built-in validation checks that prevent the uploading of several types of erroneous or impossible data points. However, data were typically first uploaded to a test database with the same validation checks so that any remaining errors were corrected prior to uploading data to Babase. The source code for Babase is freely available and downloadable²⁶.

Usage Notes

The social group history file²² provides metadata that are necessary to understand how social groups have changed over time via fissions and fusions and should be referenced when analyzing group-level data.

The data on activity budgets and diet²² were recorded during a limited number of hours each day: focal animal samples were typically only conducted between 0700 and 1100 or between 1400 and 1800, although the daily time periods for focal animal sampling have varied somewhat over the past 40 years. This means that our activity data are representative of baboon behavior during these time periods. In addition, when conditions were not favorable to focal animal sampling (e.g., the group was walking quickly or spread over a very wide area) we missed many more point samples (i.e., recordings of behavior on the minute, within focal samples) than when the animals were resting or close to each other, and these missed point samples may result in bias in which activities were recorded in some circumstances³³. For instance, in this population the amount of focal data recorded by observers is negatively correlated with the proportion of time individuals spend walking. This correlation is likely caused by frequent movement making it more difficult for observers to maintain visual contact with focal individuals, leading to more out of sight focals or interruptions in sampling. These potential sources of bias in the animals' activities and diet composition should be taken into consideration when interpreting the activity budget data.

The data on rainfall and temperature²² were collected from different weather stations, which reflect changes in the devices used to collect the data or the location of the station. It is likely that the weather readings reflect some systematic variation as a function of weather station identity. We include a unique identifier for the weather station that collected each weather data record published here. Temperature data collected by one weather station (ID: BC6) in 2022 and 2023 are not published here because during this period, the station's analog min-max thermometer was replaced with a digital minimum-maximum thermometer that produced anomalously high readings. In addition to variation as a function of weather station, the study groups have ranged over a wide and shifting area and we suspect that the study area is characterized by spatial heterogeneity in precipitation at fine temporal scales. This spatial heterogeneity means that our rainfall data provide a coarse but not a fine-grained measure of rainfall experienced by the study animals. For example, on a given day, it is possible that some rain fell within the home range of one study group, but not within the home range of another study group.

In addition to the data sets published here, the ABRP has collected and curated other data sets on the Amboseli baboon population (e.g., individual life histories, hormonal and genetic data, etc.). Many data sets included in published work in the past two decades are already available through publicly available repositories (e.g., Dryad, NCBI's Short Read Archive, MoveBank) or in the supplementary materials associated with our publications. We refer interested researchers to the original publications, where the location of cleaned and curated data sets are indicated. Researchers interested in accessing unpublished data are welcome to contact the directors of the ABRP to discuss potential collaborations, including a description of the proposed project and details on the data types of interest.

Data availability

The data²² are accessible on Dryad (<https://doi.org/10.5061/dryad.wwpzgmzf>).

Code availability

R³⁴ code used to generate figures is available on GitHub (https://github.com/CASouthworth/ABRP-Scientific_Data) and utilizes R packages *tidyverse*, *ggpubr*, *data.table*, *cowplot*, *ggarchery* and *ggnewscale*^{35–40}.

Received: 8 October 2025; Accepted: 27 January 2026;

Published online: 02 March 2026

References

- Clutton-Brock, T. & Sheldon, B. C. Individuals and populations: the role of long-term, individual-based studies of animals in ecology and evolutionary biology. *Trends Ecol. Evol.* **25**, 562–573, <https://doi.org/10.1016/j.tree.2010.08.002> (2010).
- Zipple, M. N. & Lange, E. C. How long-term studies reveal otherwise hidden phenomena: The Long-Term Animal Research Seminar Series. *Evol. Anthropol. Issues News Rev.* **30**, 108–112, <https://doi.org/10.1002/evan.21888> (2021).
- Sheldon, B. C., Kruuk, L. E. B. & Alberts, S. C. The expanding value of long-term studies of individuals in the wild. *Nat. Ecol. Evol.* **6**, 1799–1801, <https://doi.org/10.1038/s41559-022-01940-7> (2022).
- Alberts, S. C. & Altmann, J. The Amboseli Baboon Research Project: 40 Years of Continuity and Change. in *Long-Term Field Studies of Primates* (eds Kappeler, P. M. & Watts, D. P.) 261–287, https://doi.org/10.1007/978-3-642-22514-7_12 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012).
- Alberts, S. C. Social influences on survival and reproduction: Insights from a long-term study of wild baboons. *J. Anim. Ecol.* **88**, 47–66, <https://doi.org/10.1111/1365-2656.12887> (2019).
- Tung, J., Lange, E. C., Alberts, S. C. & Archie, E. A. Social and early life determinants of survival from cradle to grave: A case study in wild baboons. *Neurosci. Biobehav. Rev.* **152**, 105282, <https://doi.org/10.1016/j.neubiorev.2023.105282> (2023).
- Kingdon, J. *The Kingdon Field Guide to African Mammals*. (Princeton University Press, Princeton, NJ, 1997).
- Alberts, S. C. & Altmann, J. Balancing costs and opportunities: dispersal in male baboons. *Am. Nat.* **145**, 279–306, <https://doi.org/10.1086/285740> (1995).
- Altmann, J. & Alberts, S. C. Variability in reproductive success viewed from a life-history perspective in baboons. *Am. J. Hum. Biol.* **15**, 401–409, <https://doi.org/10.1002/ajhb.10157> (2003).
- Silk, J. B., Altmann, J. & Alberts, S. C. Social relationships among adult female baboons (*Papio cynocephalus*) I. Variation in the strength of social bonds. *Behav. Ecol. Sociobiol.* **61**, 183–195, <https://doi.org/10.1007/s00265-006-0249-2> (2006).
- Silk, J. B., Alberts, S. C. & Altmann, J. Social relationships among adult female baboons (*Papio cynocephalus*) II. Variation in the quality and stability of social bonds. *Behav. Ecol. Sociobiol.* **61**, 197–204, <https://doi.org/10.1007/s00265-006-0250-9> (2006).
- Onyango, P. O., Gesquiere, L. R., Altmann, J. & Alberts, S. C. Puberty and dispersal in a wild primate population. *Horm. Behav.* **64**, 240–249, <https://doi.org/10.1016/j.yhbeh.2013.02.014> (2013).
- Alberts, S. C. *et al.* Seasonality and long term change in a savannah environment. in *Primate Seasonality: Implications for Human Evolution* (eds Brockman, J. K. & Van Schaik, C. P.), 157–196, <https://doi.org/10.1017/cbo9780511542343.007> (Cambridge University Press, 2005).
- Gesquiere, L. R. *et al.* Thyroid hormone concentrations in female baboons: Metabolic consequences of living in a highly seasonal environment. *Horm. Behav.* **161**, 105505, <https://doi.org/10.1016/j.yhbeh.2024.105505> (2024).

15. Post, D. G. Activity patterns of yellow baboons (*Papio cynocephalus*) in the Amboseli National Park, Kenya. *Anim. Behav.* **29**, 357–374, [https://doi.org/10.1016/s0003-3472\(81\)80095-4](https://doi.org/10.1016/s0003-3472(81)80095-4) (1981).
16. Bronikowski, A. M. & Altmann, J. Foraging in a variable environment: weather patterns and the behavioral ecology of baboons. *Behav. Ecol. Sociobiol.* **39**, 11–25, <https://doi.org/10.1007/s002650050262> (1996).
17. Altmann, S. A. *Foraging for Survival: Yearling Baboons in Africa*. (University of Chicago Press, 1998).
18. Altmann, S. A. Diets of yearling female primates (*Papio cynocephalus*) predict lifetime fitness. *Proc. Natl. Acad. Sci.* **88**, 420–423, <https://doi.org/10.1073/pnas.88.2.420> (1991).
19. Altmann, J., Schoeller, D., Altmann, S. A., Muruthi, P. & Sapolsky, R. M. Body size and fatness of free-living baboons reflect food availability and activity levels. *Am. J. Primatol.* **30**, 149–161, <https://doi.org/10.1002/ajp.1350300207> (1993).
20. Charpentier, M. J. E., Tung, J., Altmann, J. & Alberts, S. C. Age at maturity in wild baboons: genetic, environmental and demographic influences. *Mol. Ecol.* **17**, 2026–2040, <https://doi.org/10.1111/j.1365-294x.2008.03724.x> (2008).
21. Tung, J., Archie, E. A., Altmann, J. & Alberts, S. C. Cumulative early life adversity predicts longevity in wild baboons. *Nat. Commun.* **7**, 11181, <https://doi.org/10.1038/ncomms11181> (2016).
22. Southworth, C. A. *et al.* Demographic, behavioral, and ecological data from a long-term field study of wild baboons in Amboseli, Kenya. *Dryad*. <https://doi.org/10.5061/dryad.wwpzgmzsf> (2025).
23. Alberts, S. C. & Altmann, J. Immigration and hybridization patterns of yellow and anubis baboons in and around Amboseli, Kenya. *Am. J. Primatol.* **53**, 139–154, <https://doi.org/10.1002/ajp.1> (2001).
24. Vilgalys, T. P. *et al.* Selection against admixture and gene regulatory divergence in a long-term primate field study. *Science* **377**, 635–641, <https://doi.org/10.1126/science.abm4917> (2022).
25. Alberts, S. C., Archie, E. A., Altmann, J. & Tung, J. Monitoring Guide for the Amboseli Baboon Research Project. https://amboselibaboons.nd.edu/assets/384683/abrp_monitoring_guide_9april2020.pdf (2020).
26. Pinc, K. O. *et al.* Babase: Technical Specifications for the Amboseli Baboon Project Data Management System. https://papio.biology.duke.edu/babase_system.html (2025).
27. Altmann, J., Altmann, S. & Hausfater, G. Physical maturation and age estimates of yellow baboons, *Papio cynocephalus*, in Amboseli National Park, Kenya. *Am. J. Primatol.* **1**, 389–399, <https://doi.org/10.1002/ajp.1350010404> (1981).
28. Alberts, S. C. & Altmann, J. Preparation and activation: determinants of age at reproductive maturity in male baboons. *Behav. Ecol. Sociobiol.* **36**, 397–406, <https://doi.org/10.1007/bf00177335> (1995).
29. Altmann, J. Observational study of behavior: sampling methods. *Behaviour* **49**, 227–266, <https://doi.org/10.1163/156853974x00534> (1974).
30. McDonald, M. & Johnson, S. ‘There’s an app for that’: a new program for the collection of behavioural field data. *Anim. Behav.* **95**, 81–87, <https://doi.org/10.1016/j.anbehav.2014.06.009> (2014).
31. PostgreSQL Global Development Group. PostgreSQL.
32. The PHP Development Team. PHP.
33. Creighton, M. J. A. *et al.* *bioRxiv*, 2025–11, <https://doi.org/10.1101/2025.11.20.689548> (2025).
34. R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, <https://doi.org/10.32614/r.manuals> (2022).
35. Wickham, H. *et al.* Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).
36. Wilke, C. O. cowplot: Streamlined Plot Theme and Plot Annotations for ‘ggplot2’ (2020).
37. Alboukadel, K. ggpubr: ‘ggplot2’ Based Publication Ready Plots (2023).
38. Dowle, M. & Srinivasan, A. data.table: Extension of ‘data.frame’. (2023).
39. Hall, M. ggarchery: Flexible Segment Geoms with Arrows for ‘ggplot2’. (2024).
40. Campitelli, E. ggnewscale: Multiple Fill and Colour Scales in ‘ggplot2’. (2025).

Acknowledgements

We thank Jeanne Altmann for her instrumental role in stewarding the Amboseli Baboon Research Project and in designing many of the ABRP’s long-term data collection protocols. In Kenya, our research was approved by the Wildlife Research Training Institute (WRTI), Kenya Wildlife Service (KWS), the National Commission for Science, Technology, and Innovation (NACOSTI), and the National Environmental Management Authority (NEMA). We also thank the University of Nairobi, the Kenya Institute of Primate Research (KIPRE), the National Museums of Kenya, the members of the Amboseli-Longido pastoralist communities, the Enduimet Wildlife Management Area, Ker & Downey Safaris, Air Kenya, and Safarilink for their cooperation and assistance in the field. We are particularly thankful to the ABRP long-term field team (R.S. Mututua, S. Sayialel, J.K. Warutere, I.L. Siodi, and L. Musembei) and to T. Wango and V. Oudu for their assistance in Nairobi. Design and programming for the ABRP database, Babase, is provided by K. Pinc. A complete set of acknowledgments of funding sources, logistical assistance, and data collection and management is available at <http://amboselibaboons.nd.edu/acknowledgements/>. We gratefully acknowledge the support of the National Science Foundation and the National Institutes of Health for the majority of the data represented here, currently through R01AG071684, R01AG075914, and R61AG078470. Current support for field-based data collection also comes from the Max Planck Institute for Evolutionary Anthropology, and we thank Duke University, Princeton University, and the University of Notre Dame for financial and logistical support.

Author contributions

C.A.S., J.C.W., S.C.A., E.A.A. and J.T. designed the research. S.C.A., E.A.A. and J.T. produced the data and provided project administration. J.B.G., N.H.L. and W.A.W. managed and curated the data. C.A.S., J.C.W., J.B.G., N.H.L., W.A.W., C.A., G.A., M.A., C.R.C., M.N.C., M.J.A.C., C.M.C., R.D., C.C.D., P.F., E.M.G., L.R.G., S.H., L.H., E.E.J., I.G.K., B.A.L., L.N., I.R.-G., B.T. and S.E.W. analyzed and visualized the data. C.A.S., J.C.W., S.C.A., E.A.A. and J.T. wrote the manuscript and all authors participated in revisions.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-026-06741-2>.

Correspondence and requests for materials should be addressed to C.A.S. or J.C.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026