

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES

- » Reproductive biology
- » Data publication and archiving
- » Biological anthropology

Life history profiles for 27 strepsirrhine primate taxa generated using captive data from the Duke Lemur Center

Sarah M. Zehr^{1,2}, Richard G. Roach¹, David Haring¹, Julie Taylor¹, Freda H. Cameron¹ and Anne D. Yoder^{1,3}

Received: 23 April 2014

Accepted: 18 June 2014

Published: 22 July 2014

Since its establishment in 1966, the Duke Lemur Center (DLC) has accumulated detailed records for nearly 4,200 individuals from over 40 strepsirrhine primate taxa—the lemurs, lorises, and galagos. Here we present verified data for 3,627 individuals of 27 taxa in the form of a life history table containing summarized species values for variables relating to ancestry, reproduction, longevity, and body mass, as well as the two raw data files containing direct and calculated variables from which this summary table is built. Large sample sizes, longitudinal data that in many cases span an animal's entire life, exact dates of events, and large numbers of individuals from closely related yet biologically diverse primate taxa make these datasets unique. This single source for verified raw data and systematically compiled species values, particularly in combination with the availability of associated biological samples and the current live colony for research, will support future studies from an enormous spectrum of disciplines.

Design Type(s)	longitudinal animal study • observation design • demographics • data integration
Measurement Type(s)	phenotypic profiling
Technology Type(s)	phenotype characterization
Factor Type(s)	
Sample Characteristic(s)	Otolemur garnettii garnettii • Galago moholi • Cheirogaleus medius • Eulemur rubriventer • Eulemur rufus • Eulemur sanfordi • Haplemur griseus griseus • Lemur catta • Microcebus murinus • ... • multi-cellular organism

¹The Duke Lemur Center, Duke University, Durham, NC 27705, USA. ²NESCent, Duke University, Durham, NC 27705, USA. ³Department of Biology, Duke University, Durham, NC 27705, USA.

Correspondence and requests for materials should be addressed to S.M.Z. (email: sarah.zehr@duke.edu)

Background and Summary

The Duke Lemur Center (DLC) here provides public availability of a life-history resource for the world's largest collection of endangered primates. Since its establishment in 1966, the DLC has been dedicated to the study and conservation of prosimian primates—the lemurs, lorises, galagos, and tarsiers—with a special interest in the lemurs of Madagascar. The lemurs, lorises, and galagos together form the strepsirrhines, the sister clade to all other living primates (i.e., haplorhines). This report contains verified data for 3,627 individuals of the 27 strepsirrhine taxa shown in Table 1. The phylogeny, divergence times, and historical biogeography of this group are increasingly well-resolved^{1,2}, giving biologists a secure historical framework within which to ask detailed questions relating to virtually all aspects of primate genotype and phenotype. The entire DLC historic collection is represented by nearly 4,200 individuals from over 40 taxa, and is the product of nearly 50 years of captive breeding, institutional exchange, and

count	Taxon	Latin_Name	Common_Name
1	CMED	<i>Cheirogaleus medius</i>	Fat-tailed dwarf lemur
2	DMAD	<i>Daubentonia madagascariensis</i>	Aye-aye
3	EALB	<i>Eulemur albifrons</i>	White-fronted brown lemur
4	ECOL	<i>Eulemur collaris</i>	Collared brown lemur
5	ECOR	<i>Eulemur coronatus</i>	Crowned lemur
6	EFLA	<i>Eulemur flavifrons</i>	Blue-eyed black lemur
7	EFUL	<i>Eulemur fulvus</i>	Common brown lemur
8	EMAC	<i>Eulemur macaco</i>	Black lemur
9	EMON	<i>Eulemur mongoz</i>	Mongoose lemur
10	ERUB	<i>Eulemur rubriventer</i>	Red-bellied lemur
11	ERUF	<i>Eulemur rufus</i>	Red-fronted brown lemur
12	ESAN	<i>Eulemur sanfordi</i>	Sanford's brown lemur
13	EUL	<i>Eulemur</i>	Eulemur hybrid
14	GMOH	<i>Galago moholi</i>	Mohol bushbaby
15	HGG	<i>Hapalemur griseus griseus</i>	Eastern lesser bamboo lemur
16	LCAT	<i>Lemur catta</i>	Ring-tailed lemur
17	LTAR	<i>Loris tardigradus</i>	Slender loris
18	MMUR	<i>Mircocebus murinus</i>	Gray mouse lemur
19	MZAZ	<i>Mirza coquereli</i>	Northern giant mouse lemur
20	NCOU	<i>Nycticebus coucang</i>	Slow loris
21	NPYG	<i>Nycticebus pygmaeus</i>	Pygmy slow loris
22	OGG	<i>Otolemur garnettii garnettii</i>	Northern greater galago
23	PCOQ	<i>Propithecus coquereli</i>	Coquerel's sifaka
24	PPOT	<i>Perodicticus potto</i>	Potto
25	VAR	<i>Varecia</i>	Varecia hybrid
26	VRUB	<i>Varecia rubra</i>	Red ruffed lemur
27	VVV	<i>Varecia variegata variegata</i>	Black-and-white ruffed lemur

Table 1. List of taxa included in the data files, including taxonomic code used in all data files (Taxon), Latin name and common name of each taxon^{4,5}.

previously wild-caught animals obtained in collaboration with local authorities in their countries of origin. The current colony size is maintained at roughly 250 animals with 11 actively breeding species, and is the largest diverse collection of captive strepsirrhines worldwide. The endangered status of the species in the collection³, combined with regulations supporting their protection in the wild, make it extremely unlikely that a colony of this magnitude and diversity could ever be re-created.

Details of birth, death, reproduction, and growth, along with extensive husbandry and medical records, have been documented for each animal throughout the colony history. Taxa included in this report represent four of the five extant lemur families, and nine of 14 recognized lemur genera^{4,5}, yielding good coverage of the Lemuroidea; representatives from the galago and loris lineages provide outgroups for the study of lemurs, and round out coverage of the strepsirrhine clade as shown in Figure 1. It is a remarkable collection of animal-associated data for a phylogenetically and biologically diverse assemblage of endangered primates. Until now, however, use of this information was both challenging and time-consuming because the data were not available in a uniform, easily searchable format—a problem that existed because data have been collected and stored in ways that span the range of technologies available at different points throughout the Center's long history. We have extracted information from various incompatible formats, transferred data to usable source files, compiled them using SAS[®] software tools, and are now able to provide large amounts of colony data in flexible and analyzable formats. The workflow used to generate the output presented here is depicted in Figure 2.

We have generated a life history table that provides species data and summary statistics for variables relating to adult, young-adult, and neonatal body mass, birth and breeding season, litter size, age at reproduction, longevity, infant mortality, activity pattern, and numbers of live individuals who are available for study in our current colony or for which biological samples have been banked. The life history table was generated with values from the two accompanying data files: the DLC Animal List, containing single-copy variables for each of the 3,627 individuals included in this paper, and the DLC Weight File, containing 65,692 weight measurements from 2,174 of those individuals over time.

The data, used independently or when combined with the research accessibility of the DLC's current live colony and the availability of affiliated biological samples for 1,012 of the individuals included in this paper, will support a large number of research projects across a diverse span of biological disciplines. As additional data from the historic colony are verified and data for newly arrived individuals obtained, future updates containing increasing numbers of variables, individuals and taxa will be made available.

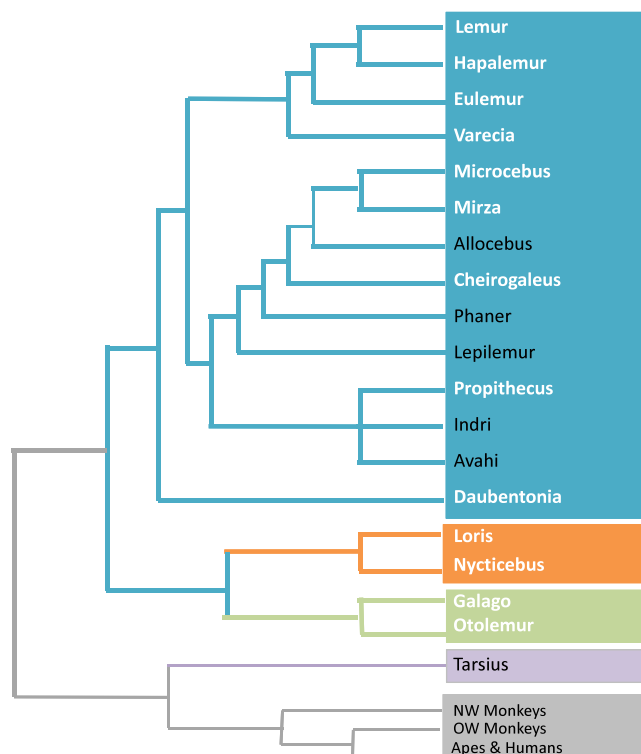


Figure 1. Primate phylogeny highlighting strepsirrhine genera included in this study. The recognized strepsirrhine genera and their phylogenetic relationships are shown^{1,2,5,6}. Genera for which data are presented are shown in white, and include nine of 14 lemur genera, along with two genera from each of the loris and galago families.

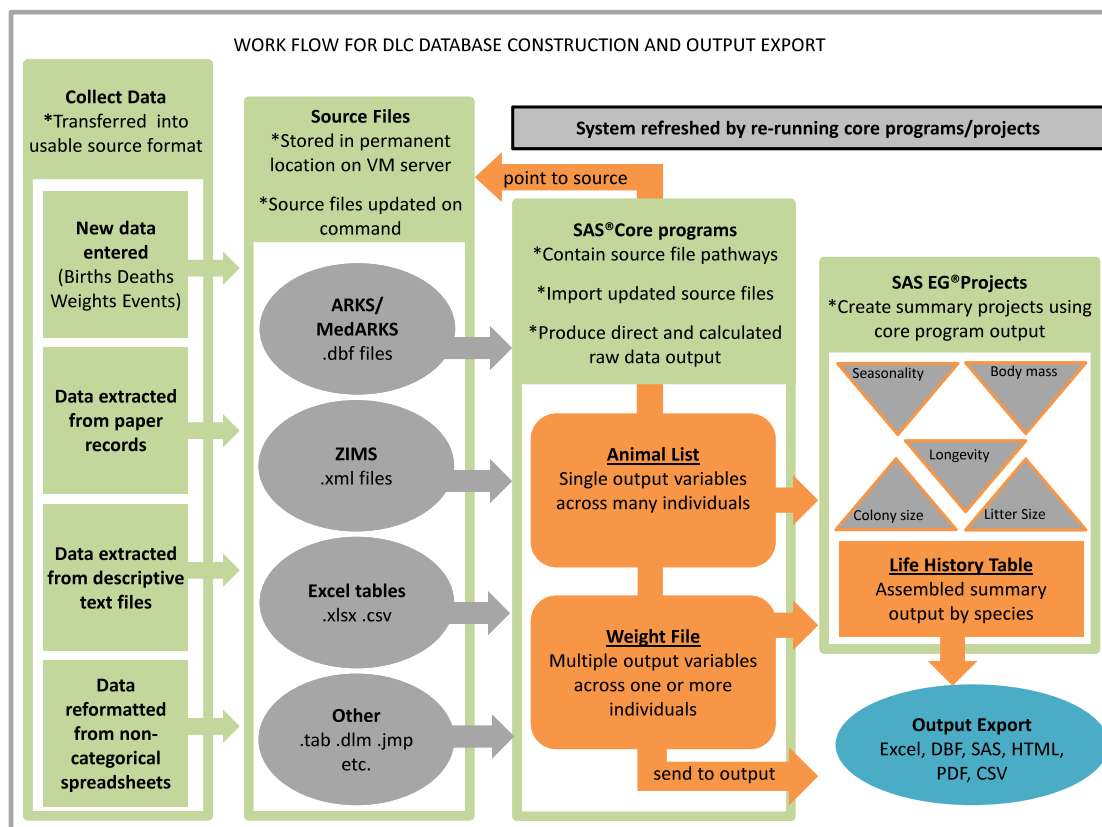


Figure 2. Workflow for DLC Database Construction. Output files contained in this database release are shown in orange. We can automatically refresh the output to include any newly entered data simply by re-running the core programs and secondary projects. By saving copies of the source files from any given date of update, older versions can be re-created by placing them in the source folder to which the core programs point. Data files presented in this paper are available in .csv format, but other possible output formats are indicated in the 'Output Export' cell.

Methods

Data source

The Duke Lemur Center is situated on 80 acres in Duke Forest, Durham NC. Individual animal records have been kept by DLC staff for a total of 4,189 primates owned by the DLC and/or housed at the DLC. Many of the diurnal animals are housed in semi-free-ranging outdoor enclosures that encourage more naturalistic behavior and social interaction. Over the course of the colony's history, 202 animals were brought in as wild-caught founders. A total of 3,229 animals have been born at the DLC, with an additional 197 individuals DLC-owned but captive born elsewhere (born to DLC-owned dams on loan at other institutions). The remainder is comprised of animals (both wild-born and captive-born) that have transferred into the colony as a loan, donation, or trade from another institution. It is an active breeding colony, so number of animals recorded is continually increasing. Data have been verified in all categories for 3,627 of these individuals, and only those are included in this report.

Data collection and entry

Animal data have been collected and entered by DLC staff according to standard operating procedures and USDA, AZA, and IACUC guidelines throughout the history of the center (United States Department of Agriculture, Association of Zoos and Aquariums, Institutional Animal Care and Use Committee respectively). Births, deaths, weights, enclosure moves, behaviors, and other significant events are recorded daily by animal care, veterinary, and research staff and subsequently entered into the permanent records by the DLC Registrar. Oldest records are mainly in hand-written or typed paper format, with a move to computerized files as that technology became available. In the mid 1990s, we introduced the use of two databases that provided the ability to link our information with that of other captive facilities through the International Species Information System (ISIS). These databases, the Animal Record Keeping System (ARKSTM; ISIS, version 4.0, 2010) and the medical version of ARKS (MedARKSTM; ISIS, version 5.54.d, 2011) are still currently in use for data entry, but do not produce analyzable output.

We are in the process of migrating to the ISIS-implemented replacement for ARKS (the Zoological Information Management System, ZIMSTM; ISIS, version 2.0, 2014) for entry and storage of husbandry data, which will allow us to maintain critical links to ISIS. Additional categories of data not supported by these databases have been entered into spreadsheet format, and the extraction of categorical data from old descriptive text records for transfer to one of the aforementioned storage formats is ongoing. Thus, future updates will contain not only newly collected colony information, but newly extracted and verified information from old and descriptive records as well, and will expand the number of individuals, taxa, and variables available.

Data extraction, importation, and compilation

The DLC database was built using SAS[®] software. Data were imported into SAS Enterprise Guide[®] [SAS, version 4.3 (2010), 5.1(2011)] from the following sources: a) ARKS and MedARKS as.dbf files b) ZIMS as.xml files and c) various categorical data types as Excel files (.csv or.xlsx format). Programs were written in SAS[®] [SAS, version 9.2 (2009), 9.3 (2011)] to extract, match, and/or join direct variables from the various source files, to perform additional calculations to create new variables, and to format data output. Additional calculations and formatting were carried out using drag and drop tools within SAS Enterprise Guide Projects. Data matching for individual animals is based on a unique DLC identifier (variable name = DLC_ID), while species-related variables are matched based on taxonomic name (variable name = taxon).

Construction of strepsirrhine life history table

The DLC Life History Table (Table 2 (available online only)) is constructed entirely from variables available in the two associated raw data files provided in this data paper (DLC Animal List and DLC Weight File). Each variable presented in the life history table is named using five terms in the format of Category_Measurement_Group_Variable_Units. The first term is the identifier to indicate a data category as follows: variables relating to sample size or animal counts have an S. Variables relating to reproduction begin with an R. Those relating to body mass have an M, and those relating to longevity and mortality an L. Finally, any variables not relating to those categories begin with an O (other). The second term in the name is the type of measurement, for example N, Mean, Max, Min, Peak or Pct. The third term identifies the group the variable is assessed for: all individuals (All), males only (M) females only (F), individuals of undetermined sex (ND), female parent (Dam), or male parent (Sire). The grouping term is omitted if inappropriate, as is the case for some variables relating to litter size, birth, and breeding season. These three qualifiers are followed by the variable core (e.g., AdultWeight, LitterSize, AgeAtDeath, etc.). Finally a unit of measure is added to the end if needed (y = years, day = days, g = grams). All life history table variable definitions are shown in Table 3 (available online only), and justification and explanation of calculations are given below.

Decisions regarding which individuals and data to include in each summary calculation were guided by our intimate knowledge of captive management and breeding practices, some of which could make certain subsets of data unreliable, and we urge users to consider these limitations as described; additional cautions are given in the 'Usage Notes' section below. Should users opt to implement different strategies for generating species summaries, however, they need only to refer to the accompanying data files for the source information. In addition to the reference table provided in this data descriptor, an analytic version of the table is available for direct use in statistical software (see 'Data Records' below).

Category S: Animal number/sample size variables: Numbers of animals in various categories were counted to provide sample sizes for specific summary variables presented in the life history table, as well as to provide criteria that researchers can use to determine which species and/or subsets of data may be appropriate for use in other projects based on sample size requirements. Number of individuals in the historic DLC colony includes animals born at the DLC, wild-born animals, animals from other institutions that transferred into the colony at any time in the DLC history, and DLC-owned animals at other institutions (i.e., all animals for which we have data). Numbers of animals in the current DLC colony reflect animals currently living on site at the DLC who are potentially available for research use. Male and female individuals are sexed at birth or acquisition, and in cases where infants were stillborn or died very young and not sexed, sex is designated as ND (not determined). Captive-born (CB) animals were born at the DLC or at another captive facility and have known dates of birth. Wild-born (WB) animals were imported by the DLC or by another institution from the animal's country of origin and have estimated dates of birth. For some individuals, origin is unknown (U) and they too have estimated dates of birth.

The age of most wild-born animals and animals of unknown origin was estimated on arrival by experienced staff and based on physical appearance, tooth wear, and other morphological characteristics. If the animal's age at capture was estimated and documented, the date of birth is assigned as follows: for seasonally breeding species, the first day of the middle month of breeding season in the country of origin in the estimated year of birth is used; for non-seasonal breeders, the month and day of acquisition is used with the estimated year of birth.

In some cases an age estimate was not documented and the animal was merely described as 'adult'. In such cases, the animal was assigned the age equal to the minimum dam age at reproduction for that

species (see discussion under ‘reproductive variables’ below) and so any ages calculated for that animal are a minimum. As such, these animals are included in calculations involving the maximum of a variable (the estimated age of an animal cannot be artificially older than the true age) but not the minimum of a variable (estimated age of an animal *could* be artificially younger than the true age). Minimum dam age at conception was used in all determinations of adult age rather than using the dam value for females and the sire value for males. This decision was based on factors of captive breeding management wherein dams are more reliably bred at earliest ages to increase numbers of breeding animals in the colony, but they may be paired with more experienced sires to increase chances of breeding success, making the sire minimum age at conception potentially less reliably accurate as an indicator of adult status than dam minimum age at conception for some species with relatively few breeding sires.

Category R: Reproductive variables: Variables relating to conception (e.g., breeding season, age at conception) use an *expected* date of conception that is calculated by subtracting gestation time from infant date of birth. In an attempt to control for premature births, for which date of conception would be erroneously estimated to be earlier than the true date, infants who did not survive at least one day were excluded from calculations involving date of conception thus excluding any infants who were stillborn due to prematurity.

Expected gestation time is assigned for each species based on DLC cases where copulation was observed and documented and the number of days to subsequent infant birth counted. *Gestation time ranges* take into account both these observations and reports of gestation time in wild populations^{5–11}. Because much of the information relating to breeding events is still embedded in descriptive records, gestation time variables are a summary rather than a true calculation, with ‘Expected_Gestation_day’ identified as the most commonly used at the DLC for each species, which in most cases lies on the lower end of the gestation range. More thorough analyses of descriptive DLC breeding records may result in fine-tuning of these values in the future, but we have confidence that these are valid estimates because their use with observed breeding behaviors in the current colony predicts birth to within a few days.

Each taxon is characterized as a seasonal (S) or a non-seasonal (NS) breeding species. Birth and Breeding season peaks are based only on infants born at the DLC with known dates of birth and are only calculated for seasonally breeding species. Non-seasonally breeding species show a ‘0’ for these variables. As discussed above, date of conception is calculated by subtracting gestation time from date of birth; date of birth is an exact value, but date of conception is an estimate. We therefore assess seasonal values using month, rather than day, of the event. Peak birth/breeding month is calculated by identifying the month in which the most events occurred for each species; an event is defined here as the birth or conception of a litter, not an individual. Peak birth/breeding season includes sequential months on either side where a) at least one third as many events took place or b) at least 20% of total number of events took place. These constraints were imposed to identify the peak, not necessarily the entire, breeding season. Importantly, their implementation systematically eliminates small tails that may artificially lengthen the true breeding season, especially for species where overall number of births was low and a single early or late pregnancy could extend it.

Minimum ages at reproduction were calculated using only individuals with known dates of birth (i.e., no wild-born animals or animals of unknown origin). Maximum ages at reproduction may include individuals with estimated dates of birth (see discussion above). Litter size variables are based only on animals born at the DLC but do include *all* DLC births (including infants that did not survive 1 day), and include the average, most common, maximum and minimum litter sizes observed for each species, as well as the frequency of the most common litter size. Birth sex ratio of male to female DLC births is calculated for each taxon. Only DLC births are used because animals brought in from the wild or from other institutions may be selected based on programmatic needs (e.g., wild founders were typically imported as male-female pairs) and may mask the underlying birth sex ratio.

Category L: Longevity variables: Maximum age is determined by the age of the oldest individual recorded, living or dead, and includes animals with estimated dates of birth as described above. Longevity was assessed using a proportional hazards model after exclusion of young infant deaths (death prior to 30 days of age) and censoring of living animals and those with uncertain status. Median longevity was derived from the average age of the nearest uncensored values above and below 50% survivorship after each was weighted by distance from the 0.5 midpoint [$((\text{Upper age} - (\text{Upper age} * \text{distance from midpoint})) + ((\text{Lower age} + (\text{Lower age} * \text{distance from midpoint})))) / 2$]. Infant mortality percentage is calculated as the percentage of infants born at the DLC who died at less than 30 days of age. Infant mortality here *does* include stillbirths. More thorough analyses of DLC descriptive records will eventually allow us to differentiate between rates of stillbirth and rates of live infant mortality.

Category M: Body mass variables: To eliminate artificially low or high values from unviable stillborn individuals, mean, maximum and minimum neonatal weight calculations include only individuals that survived at least 1 day. Neonatal body mass variables include measurements taken on day 0 (day of birth) or day 1. If both have been recorded for a single individual, the average is used to represent that individual. Inclusion of day 1 weights dramatically increases sample sizes for some species, and individuals variably gain, maintain or lose weight on day 1, so we feel that inclusion of both

measurements will yield the most accurate species values. Species values are the mean, maximum, and minimum weights across all individuals of that species.

Adult body mass calculations include all weights obtained after an individual reached *twice* the minimum dam age at reproduction for the species. This age cutoff was enforced in order to ensure that adult body mass values are not artificially lowered due to the inclusion of weights taken during the late 'near-adult body size' growth period. An additional category of young-adult body mass was created and includes weights taken when an animal was between minimum age at reproduction and twice the minimum age at reproduction. A comparison of the two categories reveals that young adult averages are indeed lower than the adult averages for 24 of 27 species, indicating that animals of most species are still growing, albeit slowly, during this period. Finally, in order to ensure that weights from wild-caught juveniles of non-exact age did not affect age-based analyses, weights from wild-caught animals at age estimates younger than dam age at first reproduction were excluded.

Mean adult and young adult weights are calculated as follows: weights taken within 60 days of death were excluded, as were weights from pregnant females. If multiple weights were obtained for an individual in a single month, those weights were averaged so as not to bias results from periods of frequent weighing (e.g., for research projects or due to health concerns). The average body weight for each individual was then generated using the mean across all monthly averages for that animal. A mean for all individuals of a species was then calculated from these individual averages. Animals whose average weight was more than two standard deviations above the mean were considered obese (adult = 46 of 1,358 individuals; young adult = 32 of 932 individuals), and removed from final analyses so as not to skew species means upward^{12,13}. The remaining (non-obese) individual means were averaged by species, and by sex where indicated. Maximum and minimum body mass values were obtained from the highest and lowest of these individual averages.

Category O: Other variables: Activity type—nocturnal or diurnal—is indicated for each species (N and D respectively). Numbers of individuals for which biological samples are available in the DLC collection is also shown for each taxon. Biological samples are opportunistically banked, and types of samples available include blood, serum, cadavers, ultra-cold tissues from major and obscure organs, RNA-later infused tissues, and formalin fixed samples. The number and type of samples available for any individual is wildly variable, with some individuals having all of the above sample types available, and others represented by as little as a single eyeball. See below under 'usage notes/additional data....' for details of how to obtain more information about samples specific to a species, individual, or sample type.

Construction of the DLC animal list and weight files

Variables included in the Animal list are single copy variables with a row entry for each of the individuals in the historic DLC colony. The weight file contains multiple copy variables with a row entry for each individual weight measurement recorded. In most cases, the core variables in these files were already being tracked in one of the source systems for current data (e.g., ARKS), and the source was populated with older data extracted from other, inaccessible, formats. Once in categorical formats, the data were compiled as described above. Additional variables were then calculated within the programs that generate the Animal List and Weight File using those core variables of directly entered data. There are core and calculated variables that are overlapping in both files that are included for ease of analysis. Variable definitions and calculations for the DLC Animal List are found in Table 4 (available online only), and for the DLC Weight File in Table 5 (available online only).

Data Records

Data record 1a, 1b

The DLC Strepsirrhine Life History Summary Table contains 91 variables for 27 taxa. There are two versions of this file provided; they contain the same information but are formatted for different uses. The first, data record 1a, is a reference version, with all variables in character format in each of 91 rows, and 27 columns each referencing a different taxon. The reference version is designed to facilitate location of particular data points for particular taxon that a user may be interested in. The second, data record 1b, is an analysis version that contains a mixture of character and numeric variables each in one of 91 columns, with a row for each of the 27 species. While it is perhaps more difficult to find a specific data point in this table, this version can be imported directly into analysis software for comparison of variables across taxa. The reference version (1a) can be accessed in the HTML version of this report (Table 2 (available online only)). Data record 1b (Data Citation 1) is stored as comma separated values, and is available from the Dryad Digital Repository. Descriptions of all Life History Table variables can be found in Table 3 (available online only) of this Data Descriptor, and as an associated file in Dryad. The dataset was last updated June 6, 2014.

Data record 2

The DLC Animal List contains 32 variables for 3,627 individuals representing 27 taxa in all stages of life (see Figure 3). This file (Data Citation 1) is stored as comma separated values, and is available from the Dryad Digital Repository. Descriptions of all variables in the Animal List can be found in Table 4

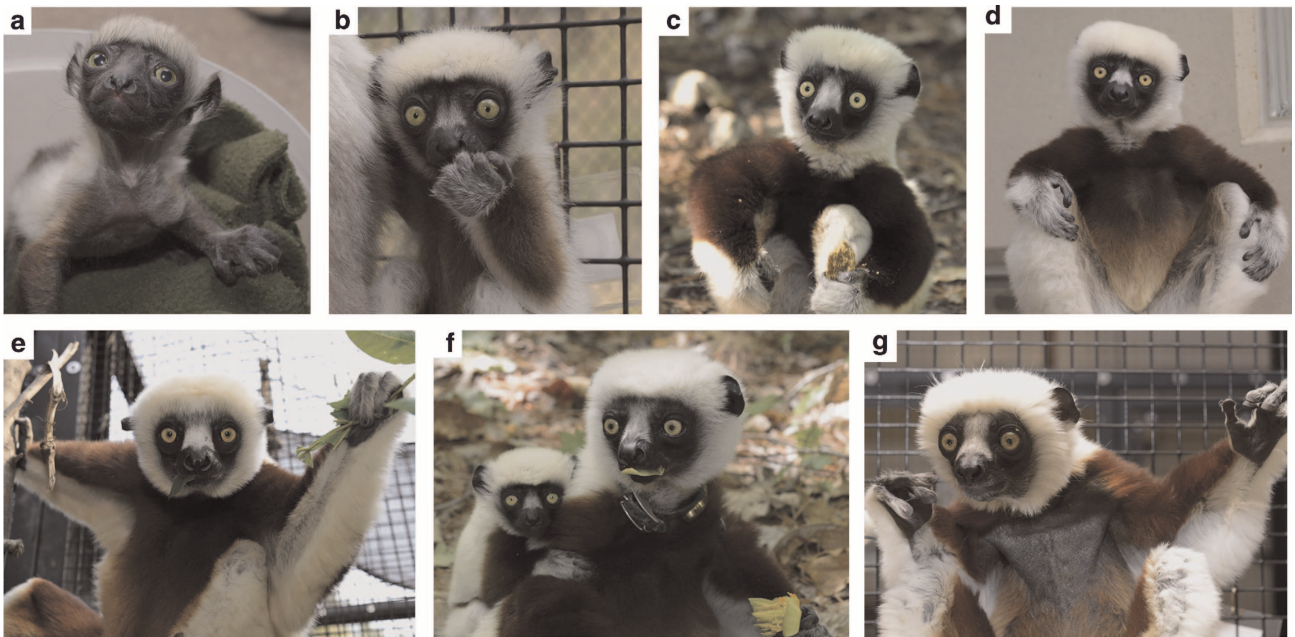


Figure 3. *Propithecus coquereli* life stages. Data files contain information for individuals at all stages of life. Shown are (a) newborn (b) infant (c) juvenile (d) yearling (e) young adult (f) reproductive adult and (g) old adult *Propithecus coquereli*.

(available online only) of this Data Descriptor, and as an associated file in Dryad. The dataset was last updated June 6, 2014.

Data record 3

The DLC Weight File contains 33 variables relating to 65,692 weight measurements across 2174 individuals and 27 taxa, and includes a large number of infant and juvenile weights (see Figure 4). This file (Data Citation 1) is stored as comma separated values, and is available from the Dryad Digital Repository. Descriptions of all variables in the Weight File can be found in Table 5 (available online only) of this Data Descriptor, and as an associated file in Dryad. The dataset was last updated June 6, 2014.

Technical Validation

Once data became accessible and analyzable in the new database, data entries were completed and verified using a variety of methods including 1) identifying missing data and scouring various DLC and ISIS records to find the information if it existed (e.g., identifying missing sires that were actually known but not entered in ARKS) 2) standardizing codes and/or text entries by conducting frequency analyses, double-checking variable types with relatively few entries and correcting typographical errors (e.g., if we find two codes have been used to flag the same type of information, a single code will be selected and the others recoded to match; if we find a misspelled taxon name that yields a single individual, the entry will be corrected), and 3) investigating impossible outliers (e.g., dates that fall outside of the project range, weights that are orders of magnitude out of the proper range, or dates of birth that yield negative calculated ages). In addition, we assessed data output from a series of known individuals to ensure that the output generated by the database programs yielded correct values in all categories. Individuals, taxa, or variables for which verification is not complete are excluded from the data presented here.

Usage Notes

Wild vs. captive populations

All data presented are from captive individuals and may not necessarily be representative of wild population values, particularly for some variables. For example, we expect that longevity in captive populations will exceed that of wild populations in most cases due to lack of predation and access to veterinary care in the former. Reproductive variables that are affected by variation in resource availability may also differ as captive populations have continual access to species-appropriate diets with no seasonal or yearly scarcity. Similarly, body mass may be higher in captive populations because they are never resource-challenged. In some cases, the degree to which captive and wild data concur will vary by species and will be affected by sample sizes and methods of analysis. We therefore warn against using these captive-derived values to *directly* assess life history variables in wild populations, but suggest that



Figure 4. Obtaining strepsirrhine infant weights. Infants are weighed on a regular basis from birth onward so that healthy growth can be monitored. The Weight File contains 14,622 weight values from 1,474 infant and juvenile individuals. Shown are (a) *Eulemur coronatus*/crowned lemur (b) *Lemur catta*/ring-tailed lemur (c) *Varecia rubra*/red ruffed lemur (d) *Eulemur flavifrons*/blue-eyed black lemur (e) *Nycticebus pygmaeus*/pygmy slow loris (f) *Propithecus coquerelii*/Coquerel's sifaka (g) *Varecia variegata variegata*/black-and-white ruffed lemur (h) *Microcebus murinus*/gray mouse lemur (i) *Daubentonia madagascariensis*/aye-aye.

researchers may be able to use them as *relative indicators* of life history variables in wild populations with caution.

For example, a comparison of body mass variables in this dataset to those for wild populations extracted from the literature for 20 lemur species¹⁴ reveals that across all species, average adult body masses at the DLC are 19% higher than averages for wild populations summarized from the literature; half of the species are fairly comparable to wild body mass (DLC animals are -1% to +11% heavier), with the other half showing greater discrepancies (DLC animals are +18% to +52% heavier). However, there are some factors that may explain some of the variability. First, in the captive data, we differentiate between adults and young adults, which would have a tendency to increase the DLC average compared to wild estimates from the literature if those contain many young adult individuals. Second, 6 of the 10 more discrepant species are flagged as having 'large variation in measurements relative to mean body size' in Table 1 of Taylor and Schwitzer¹⁴, indicating that those averages may vary seasonally, by population or by study. Finally, the most discrepant species (Eful at 52%) is so distinct from the DLC values that we

suspect the possibility that two different subspecies or quite distinct populations have been measured. If all of the cases of questionable data are excluded, the DLC animals are on average 12% heavier than wild populations. Thus, while a fine scale assessment of body size in wild populations cannot be derived from the DLC data, a comparison one species relative to another may provide insightful results.

Breeding seasonality

All breeding and birth season data are for the captive colony in North America, and are seasonally opposite of breeding and birth seasons in Madagascar, which is home to all wild lemur populations.

Timing and success of reproduction

There are some analyses that should **not** be conducted with these data because of factors associated with captive breeding management. Breeding is very strictly managed, so *individual* ages at reproduction do not give a true indication of a) individual variation in minimum, maximum, or mean age at reproduction or b) inter-birth interval. If a female does not conceive until well into adulthood, it is much more likely that she wasn't allowed to breed than that she wasn't able to breed. Because the sample sizes are large in most cases, we do feel that *species* values for the minimum and maximum ages at reproduction are reliable and as such, they are provided in the life history table.

In the DLC Animal List data file, there is a variable named 'N_Known_Offspring' that indicates the number of offspring in our records for that the individual is known to have parented. This should not be taken as a measure of relative reproductive success. In other words, dams with more known offspring are not necessarily more reproductively fit than those with fewer or no known offspring, and it may simply be that the former were given more reproductive opportunities based on management strategies. In addition, some of the offspring counts may be underestimates because a) some individuals may have had offspring at other institutions that are not accounted for in our records, and b) in cases where there are multiple possible sires of an offspring, that offspring is not counted for any of the potential sires.

Longevity

In some cases with small numbers of uncensored data points, a median longevity could not be calculated because the proportional hazards curve never drops below 50% survivorship. We see this for the aye-aye females, where the oldest female death is still young relative to other data points. As older aye-aye females die over time, these values will be able to be calculated.

Body mass: seasonal variation

The body mass summaries in the life history table are based on all weights, regardless of season in which they are taken. There is seasonal weight variation in some species, and it is particularly striking in the small nocturnal mouse and dwarf lemurs, where average summer weights (Mmur: Apr-Sept; Cmed: May-Oct) are significantly lower than the average winter weights (Mmur: Oct-March; Cmed: Nov-Apr). If the data are being used for a project that is sensitive to this, users should parcel weight data from the Weight File based on the 'MonthOfWeight' variable and use subsets accordingly.

Hybrid animals

The known *Eulemur* hybrids are a mix of between two and five of the following species: Ealb, Ecol, Eful, Eruf, Esan, and Emac. *Varecia* hybrids are a mix of Vrub and Vvv. Hybrid status is entered in ARKS and in cases where hybrid status is unknown, data output indicates a hybrid status of 'N' (not a hybrid). There are 16 animals identified taxonomically as *Eulemur* hybrids that are potentially not hybrids because at least one potential sire is of a species that matches the remainder of the animal's ancestry (DLC_ID's 5801, 5802, 5933, 5934, 6087, 5553, 5554, 1574, 2513, 2550, 2551, 3527, 3561, 1554, 6212, 2566). These animals are identifiable in the output because their taxonomic code (Eul) indicates a hybrid animal, but their hybrid status is 'N'. ZIMS *can* produce output indicating 'hybrid status unknown', so these entries will be adjusted once the migration from ARKS to ZIMS is complete.

Additional data, updates, and project information

Interested users may obtain additional access to these data files and future updated versions of data or other DLC project information as follows:

1) Direct download. As outlined above, data records are available in Dryad. Users may also download data files by connecting to the Duke Lemur Center web site (<http://lemur.duke.edu/discover/for-researchers/>) and navigating to the 'historic animal data' page where they will be guided through a brief registration (free) and then be connected to the download page. There are no costs associated with these data downloads. The download page is where future updated versions of the data described here will reside, and updates will also be deposited in Dryad on a yearly basis.

2) Specified data file requests. To inquire about or request a specific dataset that may include information not presented here, please contact corresponding author and DLC data manager, Sarah Zehr, at sarah.zehr@duke.edu. There is currently no cost for such requests, but fees may be implemented in the future to offset the cost of database maintenance.

3) Live animal or biological sample projects. For more information about use of the live DLC colony or acquisition of biological samples, please go to the Duke Lemur Center web site (<http://lemur.duke.edu>) and navigate to the research page to find more information about these classes of research, as well as contact information for the DLC Research Manager who oversees them. There are fees associated with both live animal use and biological sample purchase.

References

1. Yoder, A. D. The lemur revolution starts now: the genomic coming of age for a nonmodel organism. *Mol. Phylogenet. Evol.* **66**, 442–452 (2013).
2. Horvath, J. E. *et al.* Development and application of a phylogenetic toolkit: resolving the evolutionary history of Madagascar's lemurs. *Genome Res.* **18**, 489–499 (2008).
3. IUCN. *IUCN Red List of Threatened Species*. <http://www.iucnredlist.org> (2013).
4. Groves, C. P. *Primate Taxonomy* (Smithsonian Institution Press, 2001).
5. Mittermeier, R. A. *et al.* *Lemurs of Madagascar* 3rd edn (Conservation International, 2010).
6. Nowak, R. *Walker's Mammals of the World* 6th edn (The Johns Hopkins University Press, 1999).
7. Charles-Dominique, P. *Ecology and Behaviour of Nocturnal Primates* (Gerald Duckworth & Co. Ltd, 1977).
8. Nash, L. & Harcourt, C. Social organization of galagos in Kenyan coastal forests: *Galago garnettii*. *Am. J. Primatol.* **10**, 357–369 (1986).
9. Richard, A. In *Primate Societies: Malagasy Prosimians: Female Dominance* (University of Chicago Press, 1987).
10. Pullen, S., Bearder, S. & Dixon, A. Preliminary observations on sexual behavior and the mating system in free-ranging lesser galagos (*Galago moholi*). *Am. J. Primatol.* **51**, 79–88 (2000).
11. Quinn, A. & Wilson, D. Daubentonia madagascariensis. *Mamm. Species* **710**, 1–6 (2004).
12. Kemnitz, J. W., Guy, R. W., Flitsch, R. J., Lomiller, J. J. & Robinson, J. A. Obesity in male and female rhesus monkeys: Fat distribution, glucoregulation, and serum androgen levels. *J. Clin. Endocr. Metab.* **69**, 287–293 (1989).
13. Terranova, C. J. & Coffman, B. S. Body weights of wild and captive lemurs. *Zoo Biol.* **16**, 17–30 (1997).
14. Taylor, L. A. & Schwitzer, C. Body masses of wild lemurs. *Lemur News* **16**, 34–40 (2011/12).

Data Citation

1. Zehr, S. M. *et al.* Dryad <http://dx.doi.org/10.5061/dryad.fj974> (2014).

Acknowledgements

The project was supported by the National Science Foundation (DBI 1258440), the National Evolutionary Synthesis Center, the Duke Lemur Center, and Duke University Natural Sciences; the Duke Lemur Center is supported by the National Science Foundation (DBI 1258440), Duke University, and private donations. We thank M. Dean, E. Larson, B. Deahl, K. Skrinak, P. Daniels for IT support, K. Cranston and T. Vision for helpful comments and discussion, and the dedicated DLC animal care staff, past and present, without whom there would be no data. This is DLC publication number 1271. Dedicated to the memory of Richard G. Roach.

Author Contributions

S.M.Z. designed the project, programmed calculated variables and import of ZIMS data, entered and validated data, built the programs that generated the life history summary, and led the writing. R.G.R. designed and programmed the core data import systems from ARKS and MedARKS and associated calculated variables. D.H. oversaw entry and validation of husbandry data in ARKS and took all photographs. J.T. contributed to design of the Life History table and oversaw entry and validation of medical data in MedARKS. F.H.C. contributed to the design and programming of core variables. A.D.Y. is the DLC director and PI of NSF DBI 1258440 that supported the project. All authors contributed to, read and approved the manuscript.

Additional information

Tables 2–5 are only available in the online version of this paper.

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Zehr, S. M. *et al.* Life history profiles for 27 strepsirrhine primate taxa generated using captive data from the Duke Lemur Center. *Sci. Data* **1**:140019 doi: 10.1038/sdata.2014.19 (2014).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0>

Metadata associated with this Data Descriptor is available at <http://www.nature.com/sdata/> and is released under the CC0 waiver to maximize reuse.