

# SCIENTIFIC DATA

## OPEN Data Descriptor: HormoneBase, a population-level database of steroid hormone levels across vertebrates

Received: 7 December 2017

Accepted: 29 March 2018

Published: 22 May 2018

Maren N. Vitousek<sup>1,\*</sup>, Michele A. Johnson<sup>2,\*</sup>, Jeremy W. Donald<sup>3</sup>, Clinton D. Francis<sup>4</sup>, Matthew J. Fuxjager<sup>5</sup>, Wolfgang Goymann<sup>6</sup>, Michaela Hau<sup>6</sup>, Jerry F. Husak<sup>7</sup>, Bonnie K. Kircher<sup>8</sup>, Rosemary Knapp<sup>9</sup>, Lynn B. Martin<sup>10</sup>, Eliot T. Miller<sup>11</sup>, Laura A. Schoenle<sup>10</sup>, Jennifer J. Uehling<sup>1</sup> & Tony D. Williams<sup>12</sup>

Hormones are central regulators of organismal function and flexibility that mediate a diversity of phenotypic traits from early development through senescence. Yet despite these important roles, basic questions about how and why hormone systems vary within and across species remain unanswered. Here we describe HormoneBase, a database of circulating steroid hormone levels and their variation across vertebrates. This database aims to provide all available data on the mean, variation, and range of plasma glucocorticoids (both baseline and stress-induced) and androgens in free-living and un-manipulated adult vertebrates. HormoneBase ([www.HormoneBase.org](http://www.HormoneBase.org)) currently includes >6,580 entries from 476 species, reported in 648 publications from 1967 to 2015, and unpublished datasets. Entries are associated with data on the species and population, sex, year and month of study, geographic coordinates, life history stage, method and latency of hormone sampling, and analysis technique. This novel resource could be used for analyses of the function and evolution of hormone systems, and the relationships between hormonal variation and a variety of processes including phenotypic variation, fitness, and species distributions.

Design Type(s)	data integration objective • species comparison design
Measurement Type(s)	blood steroid hormone level
Technology Type(s)	data item extraction from journal article
Factor Type(s)	
Sample Characteristic(s)	Vertebrata • blood plasma

<sup>1</sup>Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA. <sup>2</sup>Department of Biology, Trinity University, San Antonio, TX 78212, USA. <sup>3</sup>Coates Library, Trinity University, San Antonio, TX 78212, USA. <sup>4</sup>Biological Sciences Department, California Polytechnic State University, San Luis Obispo, CA 93407, USA. <sup>5</sup>Department of Biology, Wake Forest University, Winston-Salem, NC 27109, USA. <sup>6</sup>Max Planck Institute for Ornithology, Seewiesen 82319, Germany. <sup>7</sup>Department of Biology, University of St. Thomas, St. Paul, MN 55105, USA. <sup>8</sup>Department of Biology, University of Florida, Gainesville, FL 32608, USA. <sup>9</sup>Department of Biology, University of Oklahoma, Norman, OK 73019, USA. <sup>10</sup>Department of Global Health, University of South Florida, Tampa, FL 33620, USA. <sup>11</sup>Cornell Lab of Ornithology, Ithaca, NY 14850, USA. <sup>12</sup>Department of Biological Sciences, Simon Fraser University, Burnaby, BC V5A 1S6, Canada. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to M.N.V. (email: [mnv6@cornell.edu](mailto:mnv6@cornell.edu)) or to M.A.J. (email: [michele.johnson@trinity.edu](mailto:michele.johnson@trinity.edu)).

## Background & Summary

Hormones are central regulators of phenotype, whose effects span multiple fields of research, from molecular biology to population biology<sup>1–5</sup>. Because of their role in regulating organismal function and flexibility, selection might be expected to constrain hormone levels or their context-dependent flexibility around one or more fitness optima<sup>6–9</sup>. Nevertheless, endocrine responses vary markedly both within and among populations<sup>8,10–12</sup>. Why do some individuals mount a hormonal response that is two or more orders of magnitude greater than others, when faced with the same stimulus? Similarly, why have some species evolved to express plasma testosterone levels that are an order of magnitude greater than others during reproduction, when testosterone mediates the same basic reproductive processes?

A particularly promising approach to answering such questions – and many others of broad interest to animal behaviour and organismal biology – lies in large-scale comparative analyses of the multitude of endocrine data that have been collected over the past several decades. Such analyses, conducted within a rigorous phylogenetic, environmental, and life-history framework, have the potential to illuminate the factors driving divergence in the hormonal mechanisms of behaviour, physiology, and morphology<sup>13,14</sup>. To date, most analyses have focused on relatively small taxonomic scales, and on comparing mean trait values across populations and species<sup>15–19</sup>. However, resources are rapidly becoming available to aggregate and analyse decades of available data on circulating hormone levels and their variation within free-living populations, across taxonomic groups. Identifying and characterizing the variation in endocrine traits, and their links with environment, life history, and fitness, could provide insight into how endocrine systems evolve, and how selection on these phenotypic integrators may influence the dynamics and distribution of populations<sup>20–23</sup>.

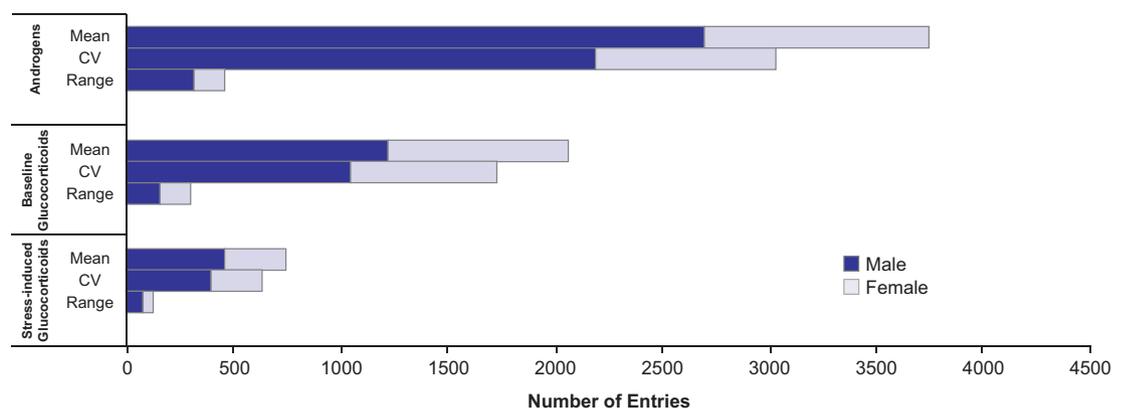
In this context, we present HormoneBase, a resource of compiled endocrine data across vertebrates. Included in this dataset are >6,580 measures of mean and within-population variation in glucocorticoids and androgens from 476 species (Figs 1,2; Table 1) that were reported in 648 publications – and additional unpublished resources – between 1967 and 2015. Additional information on geographic location (Fig. 3), life history, study design, and time period accompanies each entry. By making HormoneBase publicly available we aim to encourage data sharing across the scientific community and facilitate research into the function and evolution of physiological traits.

## Methods

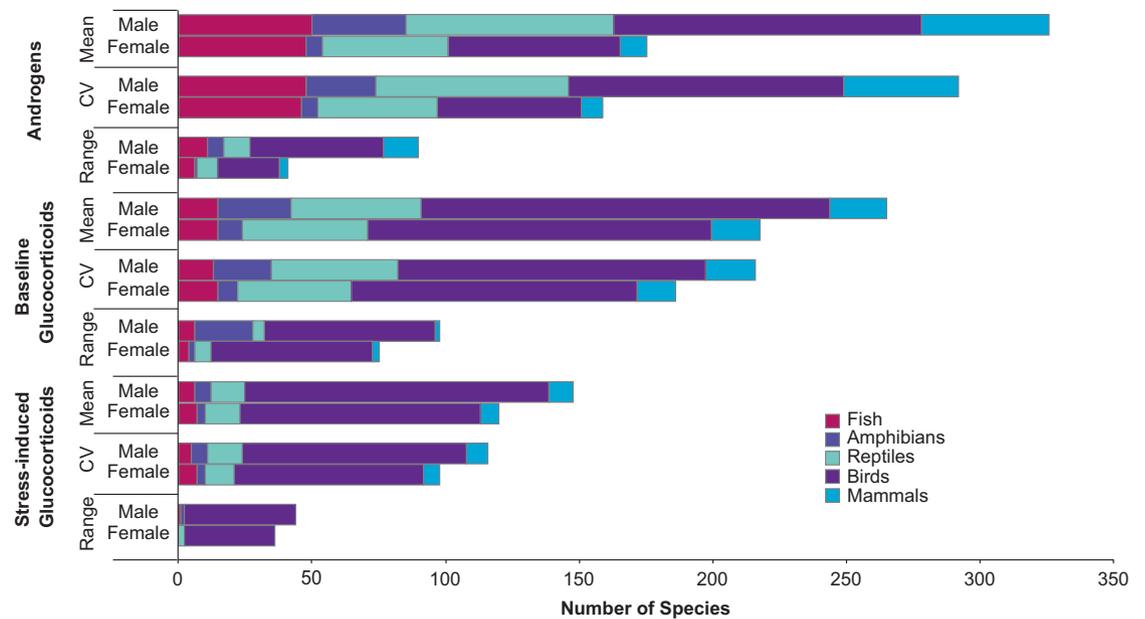
### Hormonal Data

Endocrine data were obtained from publications, and from several unpublished datasets (Data Citation 1). We searched for studies that conformed to our inclusion criteria using: (i) online academic databases (e.g., Google Scholar, Web of Science), and (ii) cross-referencing from other published works. Studies were selected for inclusion if they included data on circulating glucocorticoids (baseline or stress-induced corticosterone/cortisol) or androgens (testosterone/11-ketotestosterone) that: (i) were from free-living populations, (ii) were collected from adults that had not been subject to an experimental manipulation prior to sampling (e.g., of hormones or the environment), (iii) measured plasma levels, (iv) did not pool data across males and females, or across adults and juveniles, and (v) were reported in or could be converted to a standard unit of measurement (ng/mL).

Published values were obtained from text, tables, or supplementary materials, or extracted from published figures using the program Data Thief III (<http://datathief.org>). Entries include mean circulating concentrations (ng/mL) for each population/group and time period; whenever possible, data on within-population variation (coefficient of variation, standard error), range (maximum and



**Figure 1.** Total data entries in HormoneBase for each steroid by measurement type and sex. Within each category, counts are shown separately for mean, coefficient of variation, and range.



**Figure 2.** The number of species with data on mean hormone concentrations in HormoneBase. Counts are shown separately for males and females, and for androgens, baseline glucocorticoids, and stress-induced glucocorticoids.

	Birds		Reptiles		Amphibians		Mammals		Fish		Total	
	M	F	M	F	M	F	M	F	M	F	M	F
Androgens	1002	366	497	176	217	78	270	21	713	436	2699	1077
Baseline GCs	746	488	230	167	104	45	59	60	89	73	1228	833
Stress-induced GCs	352	212	26	28	29	7	20	27	23	20	450	294

**Table 1.** The taxonomic distribution of entries of mean circulating hormone levels in HormoneBase for males (M) and females (F).

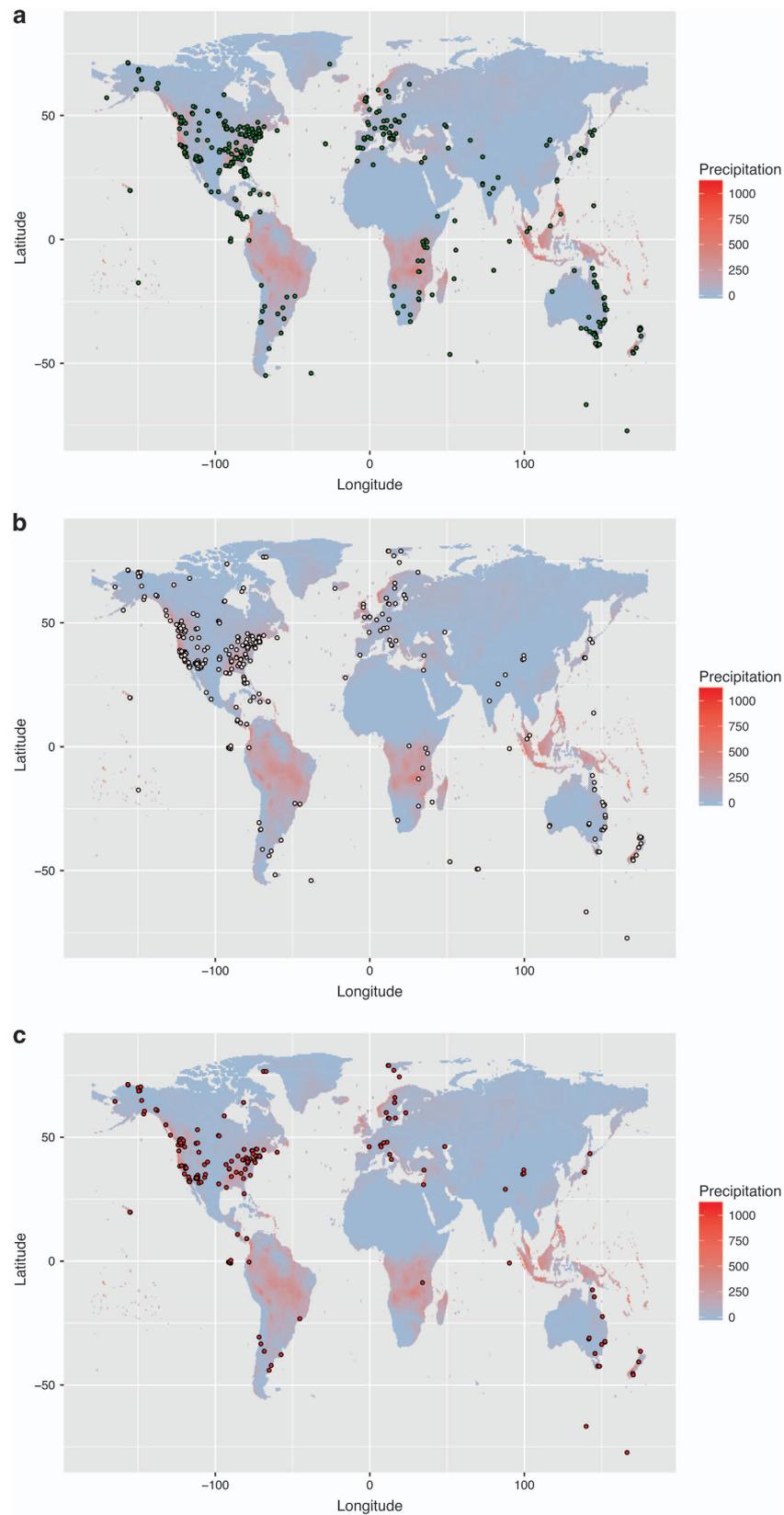
minimum values), and sample size are also included. When papers did not directly report the coefficient of variation (CV), it was calculated from the standard deviation (SD) or standard error (SE) and sample size ( $n$ ), according to the following formulas:  $CV = \frac{SD}{\text{mean}} * 100$  or  $CV = \frac{SE * \sqrt{n}}{\text{mean}} * 100$ . If papers reported that outliers had been excluded we noted this for each hormone measure, and noted the criteria for exclusion where provided.

When a single reference reported multiple means for different groups of individuals (e.g., populations or life history stages), or from different time points, data were entered on separate lines. In cases where papers reported a single hormonal mean from data collected across multiple populations, the location of up to three of the sampled populations was noted in the entry. When stress-induced glucocorticoid levels were measured at multiple time points during a standardized stress series, only the time period at which mean glucocorticoid levels were highest was included.

The decision was made to focus HormoneBase on androgens and glucocorticoids because these are currently the most widely sampled hormones across vertebrates. Because hormone concentrations are not directly comparable across biological matrices, we included only plasma hormone concentrations. Hormone levels are also increasingly being measured in other biological matrices (e.g., feces, feathers) but these sample types are not very well-suited for large-scale comparative analyses because they use hormone metabolites, which differ within and across species and assay/antibody types<sup>24</sup>.

### Sample Collection and Assay Method Data

Because sampling method and assay technique may influence circulating hormone levels, we included specific information about capture, sampling, and assay approach. The time of day (i.e., range of hours) that samples were collected was recorded as provided. The specific method used to capture free-living individuals was noted, and the capture/sampling method assigned to one of three categories. “Active” sampling methods were those in which a blood sample was obtained rapidly and within a known period



**Figure 3.** The geographical distribution of entries in HormoneBase. Points represent the location of measurements of (a) androgens, (b) baseline glucocorticoids, and (c) stress-induced glucocorticoids. Precipitation patterns reflect sums for December 2015 and were acquired from the CRU-TS 4.0 Climate Database<sup>36</sup>.

of time after targeting a previously undisturbed animal. “Passive” methods are those in which animals were sampled after an unknown period of restraint (e.g., non-continuously monitored traps or nets). “Attractant” methods are those in which animals were drawn to the site of capture using some type of attractant (e.g., song playbacks, baited traps). The maximum sampling latency (interval from capture to blood sampling) was recorded for androgens and baseline glucocorticoids, and the type of acute stressor and the interval from initial capture to the collection of a stress-induced glucocorticoid sample were also recorded.

To explore and control for potential differences in assay technique, we included information on the assay method used to assess plasma hormone levels (e.g., radioimmunoassay, enzyme immunoassay), and, where provided, the specific antibody or commercial kit that was used. Because laboratory identity can also influence measured hormone levels<sup>25</sup>, we recorded the identity of the laboratory in which hormone assays were conducted. For collaborative papers that did not identify where assays were conducted, and were the product of multiple endocrine laboratories, we arbitrarily assigned one of the collaborating laboratories as the assay location.

### Taxonomic and Geographical Data

All endocrine data include associated taxonomic information, using common and scientific names. Where relevant, scientific names were updated to reflect recent reclassifications. Taxonomy was determined using major lineage-specific trees (ray-finned fishes<sup>26</sup>, amphibians<sup>27,28</sup>, mammals<sup>29</sup>, squamates<sup>30</sup>, turtles<sup>31</sup>, and birds<sup>32</sup>).

The location name, geographic coordinates (latitude and longitude in degrees decimal), and elevation (in meters) of the population from which the data were collected are also recorded for each entry. When not provided in the original publication, approximate geographic coordinates and elevation were determined by searching for the location name in Google Earth.

### Temporal and Life History Data

To enable assessment of seasonal and life history patterns<sup>17,33,34</sup>, we included information on the time period of sampling as reported (the month and year in which data were collected) and the life history stage of sampled individuals. Measurements were characterized as coming from breeding or non-breeding individuals, or a combination of the two. Designations were based on author classifications when provided. When life history stage was not provided in the original data source, samples were classified as coming from a combination of breeding and non-breeding individuals, except in cases where seasonally breeding populations were sampled only during months that did not overlap with the breeding season.

When life history sub-stage was provided in the original data source, this information was also included in the database. To provide some standardization across species, and widely varying terminology, reported sub-stages were combined into fourteen categories: pre-breeding, courtship, incubation, copulation, gravid/pregnant, non-gravid/pregnant, laying, young care, lactation, post-breeding, migration, torpor, hibernation, pre-basic moult. When information about life history sub-stage was not contained in the original data source, this field was left blank. An associated column provides information about whether the sampled individuals were confirmed to be in a given life history stage (e.g., incubating birds captured off their nests), or whether the life history stage reflected the typical stage for individuals in that population at the time of sampling (e.g., birds sampled in mist nets during the breeding season but not traced to a specific nest). For birds, information on moult status was also recorded as provided<sup>35</sup>.

### Data Records

The HormoneBase dataset (Data Citation 1) is provided as two comma-separated values text files: a single file that includes all data described above (HormoneBase\_v1.csv), and a file that contains the reference information for the source of each entry (HormoneBase\_references\_v1.csv). Variable names are provided in the first row, with details of each variable and units measured summarized in Table 2 (available online only). These files are accompanied by a metadata pdf file (HormoneBase\_MetadataData.pdf).

### Technical Validation

The data presented in HormoneBase are primarily from published, peer-reviewed sources, but also contain unpublished data provided by authors. Data entry was initially proofed by each lab that entered the data to confirm that the entries matched reported data. Upon submission to the central repository, two members of the database entry team independently examined each entry to identify incomplete entries or extreme values. All hormone measures were also mapped onto a phylogeny to reveal putative taxonomic outliers. When such cases were identified, entries were confirmed or corrected by consulting the original source.

### Usage Notes

The data are available to access and download from *Figshare* repository (Data Citation 1). Three files are provided:

1. HormoneBase\_v1.csv
2. HormoneBase\_references\_v1.csv
3. Metadata.pdf

## References

1. Adkins-Regan, E. *Hormones and Animal Social Behavior* (Princeton University Press, 2005).
2. Lösel, R. & Wehling, M. Nongenomic actions of steroid hormones. *Nat. Rev. Mol. Cell Biol* **4**, 46–55, <https://doi.org/10.1038/nrm1009> (2003).
3. Beato, M. Gene regulation by steroid hormones. *Cell* **56**, 335–344 (1989).
4. Dantzer, B. *et al.* Density triggers maternal hormones that increase adaptive offspring growth in a wild mammal. *Science* **340**, 1215–1217, <https://doi.org/10.1126/science.1235765> (2013).
5. Duckworth, R. A., Belloni, V. & Anderson, S. R. Cycles of species replacement emerge from locally induced maternal effects on offspring behavior in a passerine bird. *Science* **347**, 875–877 (2015).
6. Hau, M. Regulation of male traits by testosterone: implications for the evolution of vertebrate life histories. *BioEssays* **29**, 133–144, <https://doi.org/10.1002/bies.20524> (2007).
7. McGlothlin, J. W. & Ketterson, E. D. Hormone-mediated suites as adaptations and evolutionary constraints. *Philos. Trans. R. Soc. B Biol. Sci* **363**, 1611–1620 <https://doi.org/10.1098/rstb.2007.0002> (2008).
8. Williams, T. D. Individual variation in endocrine systems: moving beyond the ‘tyranny of the Golden Mean’. *Philos. Trans. R. Soc. B Biol. Sci.* **363**, 1687–1698, <https://doi.org/10.1098/rstb.2007.0003> (2008).
9. Cohen, A. A., Martin, L. B., Wingfield, J. C., McWilliams, S. R. & Dunne, J. A. Physiological regulatory networks: ecological roles and evolutionary constraints. *Trends Ecol. Evol* **27**, 428–435, <https://doi.org/10.1016/j.tree.2012.04.008> (2012).
10. Cockrem, J. F. Individual variation in glucocorticoid stress responses in animals. *Gen. Comp. Endocrinol* **181**, 45–58, <https://doi.org/10.1016/j.ygcen.2012.11.025> (2013).
11. Kempnaers, B., Peters, A. & Foerster, K. Sources of individual variation in plasma testosterone levels. *Philos. Trans. R. Soc. B Biol. Sci* **363**, 1711–1723, <https://doi.org/10.1098/rstb.2007.0001> (2008).
12. Stedman, J. M., Hallinger, K. K., Winkler, D. W. & Vitousek, M. N. Heritable variation in circulating glucocorticoids and endocrine flexibility in a free-living songbird. *J. Evol. Biol* **30**, 1724–1735, <https://doi.org/10.1111/jeb.13135> (2017).
13. Jessop, T. S. *et al.* Multiscale evaluation of thermal dependence in the glucocorticoid response of vertebrates. *Am. Nat.* **188**, 342–356 (2016).
14. Crespi, E. J., Williams, T. D., Jessop, T. S. & Delehanty, B. Life history and the ecology of stress: how do glucocorticoid hormones influence life-history variation in animals? *Funct. Ecol* **27**, 93–106, <https://doi.org/10.1111/1365-2435.12009> (2013).
15. Bókony, V. *et al.* Stress response and the value of reproduction: are birds prudent parents? *Am. Nat* **173**, 589–598, <https://doi.org/10.1086/597610> (2009).
16. Eikenaar, C., Husak, J., Escallón, C. & Moore, I. T. Variation in testosterone and corticosterone in amphibians and reptiles: relationships with latitude, elevation, and breeding season length. *Am. Nat* **180**, 642–654, <https://doi.org/10.1086/667891> (2012).
17. Hau, M., Ricklefs, R. E., Wikelski, M., Lee, K. A. & Brawn, J. D. Corticosterone, testosterone and life-history strategies of birds. *Proc. R. Soc. B Biol. Sci* **277**, 3203–3212, <https://doi.org/10.1098/rspb.2010.0673> (2010).
18. Goymann, W. *et al.* Testosterone in tropical birds: effects of environmental and social factors. *Am. Nat.* **164**, 327–334 (2004).
19. Goymann, W. & Wingfield, J. C. Male-to-female testosterone ratios, dimorphism, and life history - what does it really tell us? *Behav. Ecol.* **25**, 685–699 (2014).
20. Zera, A. J., Harshman, L. G. & Williams, T. D. Evolutionary endocrinology: the developing synthesis between endocrinology and evolutionary genetics. *Annu. Rev. Ecol. Syst* **38**, 793–817, <https://doi.org/10.1146/annurev.ecolsys.38.091206.095615> (2007).
21. Wingfield, J. C. The comparative biology of environmental stress: Behavioural endocrinology and variation in ability to cope with novel, changing environments. *Anim. Behav* **85**, 1127–1133, <https://doi.org/10.1016/j.anbehav.2013.02.018> (2013).
22. Taff, C. C. & Vitousek, M. N. Endocrine flexibility: optimizing phenotypes in a dynamic world? *Trends Ecol. Evol* **31**, 476–488, <https://doi.org/10.1016/j.tree.2016.03.005> (2016).
23. Martin, L. B., Burgan, S. C., Adelman, J. S. & Gervasi, S. S. Host competence: an organismal trait to integrate immunology and epidemiology. *Integr. Comp. Biol.* **56**, 1225–1237 (2016).
24. Goymann, W. On the use of non-invasive hormone research in uncontrolled, natural environments: the problem with sex, diet, metabolic rate, and the individual. *Methods Ecol. Evol.* **3**, 757–765, <https://doi.org/10.1111/j.2041-210X.2012.00203.x> (2012).
25. Fanson, K. V., Németh, Z., Ramenofsky, M., Wingfield, J. C. & Buchanan, K. L. Inter-laboratory variation in corticosterone measurement: Implications for comparative ecological and evolutionary studies. *Methods Ecol. Evol* **8**, 1745–1754, <https://doi.org/10.1111/2041-210X.12821> (2017).
26. Rabosky, D. L. *et al.* Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat. Commun* **4**, 1–8, <https://doi.org/10.1038/ncomms2958> (2013).
27. Pyron, R. A. & Wiens, J. J. A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Mol. Phylogenet. Evol* **61**, 543–583, <https://doi.org/10.1016/j.ympev.2011.06.012> (2011).
28. Eastman, J. M., Harmon, L. J. & Tank, D. C. Congruification: support for time scaling large phylogenetic trees. *Methods Ecol. Evol* **4**, 688–691, <https://doi.org/10.1111/2041-210X.12051> (2013).
29. Bininda-Emonds, O. R. P. *et al.* The delayed rise of present-day mammals. *Nature* **446**, 507–512, <https://doi.org/10.1038/nature05634> (2007).
30. Pyron, R. A., Burbrink, F. T. & Wiens, J. J. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evol. Biol* **13**, 93, <https://doi.org/10.1186/1471-2148-13-93> (2013).
31. Jaffe, A. L., Slater, G. J. & Alfaro, M. E. The evolution of island gigantism and body size variation in tortoises and turtles. *Biol. Lett* **7**, 558–561, <https://doi.org/10.1098/rsbl.2010.1084> (2011).
32. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, O. The global diversity of birds in space and time. *Nature* **491**, 444–448, <https://doi.org/10.1038/nature11631> (2012).
33. Romero, L. M. Seasonal changes in plasma glucocorticoid concentrations in free-living vertebrates. *Gen. Comp. Endocrinol.* **128**, 1–24 (2002).
34. Wingfield, J. C. & Sapolsky, R. M. Reproduction and resistance to stress: when and how. *J. Neuroendocrinol.* **15**, 711–724 (2003).
35. Romero, L. M., Bauer, C. M., de Bruijn, R. & Lattin, C. R. Seasonal rhythms in *Stress: Neuroendocrinology and Neurobiology* 421–427, <https://doi.org/10.1016/B978-0-12-802175-0.00042-5> (Elsevier, 2017).
36. Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *Int. J. Climatol* **34**, 623–642, <https://doi.org/10.1002/joc.3711> (2014).

## Data Citation

1. Vitousek, M. N. *et al.* *figshare* <https://doi.org/10.6084/m9.figshare.5649727> (2017).

## Acknowledgements

We wish to thank Sarah Burgan, Stefania Casagrande, Pete Collins, Julia Cramer, Travis Flock, Jennifer Hoots, Michael Jessel, Lucia Mentasana, Meredith Miles, Steffanie Munguia, Samantha Murphy, Sophie Nicolich-Henkin, Sara Ocasio, Eric Schuppe, Beth Skinner, Jocelyn Stedman, Ashley Stenstrom, Sarah Talamantes, and Andrew Wang for assistance compiling hormone data. We are especially grateful to John Wingfield for submitting a substantial amount of unpublished data for inclusion in HormoneBase, and to Patrick Kelley for compiling these data.

## Author Contributions

M.N.V. and M.A.J. conceived of the idea. M.N.V., M.A.J., and J.F.H. organized the working group. All authors contributed to project design. M.N.V., M.A.J., M.F., W.G., M.H., J.F.H., B.K.K., R.K., L.B.M., L.A.S., T.D.W., and J.J.U. compiled data. JWD constructed the working relational database and provided informational support. M.A.J. and B.K.K. entered and quality checked all compiled data. C.D.F. and E.T. M. also contributed to quality checks. M.N.V. and M.A.J. drafted the paper, and all authors contributed to revising the manuscript.

## Additional Information

Table 2 is only available in the online version of this paper.

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

**How to cite this article:** Vitousek, M. N. *et al.* HormoneBase, a population-level database of steroid hormone levels across vertebrates. *Sci. Data* 5:180097 doi: 10.1038/sdata.2018.97 (2018).

**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files made available in this article.

© The Author(s) 2018