Animal chromosome counts reveal a similar range of chromosome numbers but with less polyploidy in animals compared to flowering plants

Cristian Román-Palacios1 | Cesar A. Medina2 | Shing H. Zhan1,3 | Michael S. Barker1

1Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA
2Department of Neuroscience, University of Arizona, Tucson, AZ, USA
3Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC, Canada

Correspondence
Cristian Román-Palacios, Center for Biodiversity, Temple University, Philadelphia, PA 19122, USA.
Email: cromanpa94@gmail.com

Abstract
Understanding the mechanisms that underlie chromosome evolution could provide insights into the processes underpinning the origin, persistence and evolutionary tempo of lineages. Here, we present the first database of chromosome counts for animals (the Animal Chromosome Count database, ACC) summarizing chromosome numbers for ~15,000 species. We found remarkable a similarity in the distribution of chromosome counts between animals and flowering plants. Nevertheless, the similarity in the distribution of chromosome numbers between animals and plants is likely to be explained by different drivers. For instance, we found that while animals and flowering plants exhibit similar frequencies of speciation-related changes in chromosome number, plant speciation is more often related to changes in ploidy. By leveraging the largest data set of chromosome counts for animals, we describe a previously undocumented pattern across the Tree of Life—animals and flowering plants show remarkably similar distributions of haploid chromosome numbers.

KEYWORDS
Animalia, chromosome counts, database, diploid numbers, haploid numbers, karyotype, Metazoa

1 | INTRODUCTION

The number of chromosomes in a nuclear genome is a fundamental aspect of eukaryotic biology. Differences in chromosome number and arrangement are also one of the longest studied topics in comparative genomics (e.g. Grant, 1981; King, 1995; Stebbins, 1950; White, 1973). Despite this long history, how and why chromosome numbers evolve remain poorly understood. One obstacle to improving comparative analyses and hypothesis testing is a lack of easily available chromosome counts across the Tree of Life. Although many chromosome counts are available in the published literature (Peruzzi & Bedini, 2014), these have only been made easily available for specific clades including plants (Rice et al., 2015), coleoptera (Blackmon & Demuth, 2015), polyneoptera (Sylvester & Blackmon, 2020), amphibians (Perkins et al., 2019), mammals (Blackmon et al., 2017) and fish (Arai, 2011; Martinez et al., 2015) along with other groups scattered across the Tree of Life (The Tree of Sex Consortium, 2014). Analyses of these publicly available data have driven new understandings of chromosome and genome evolution (e.g. Barker et al., 2016; Salman-Minkov et al., 2016; Zenil-Ferguson et al., 2019; Zhan et al., 2014). Increasing access to other publicly available chromosome counts from the literature will be important for understanding how and why chromosome numbers vary across the Tree of Life.

Here, we present the first database of chromosome counts across all animal clades summarizing haploid and diploid counts for
nearly 15,000 animal species across 21 phyla. We used the current release of the Animal Chromosome Count Database (ACC) to examine fundamental questions about the distribution of chromosome counts and evolution of haploid numbers in animals. These included inferring the frequency of polyploidy and estimating how often speciation is associated with changes in chromosome numbers or coupled with polyploidization. We also compared the patterns of chromosomal evolution in animals with those in angiosperms and ferns. Future studies using the ACC could examine questions related to the association between speciation and chromosomal changes in animals (e.g. Rieseberg, 2001), the importance of polyploidy in animals (e.g. Hallinan & Lindberg, 2011; Li et al., 2018) or even the causes of chromosomal variation among clades and over time (e.g. Martinez et al., 2015; Ross et al., 2015). We expect the newly compiled public animal chromosome count database to fuel new comparative analyses that further our understanding of the forces driving chromosome evolution across the Tree of Life.

2 | METHODS

2.1 | Data collection

We compiled animal chromosome counts from primary sources, books and published data sets (e.g. Benazzi & Benazzi-Lentati, 1976; Makino, 1951, along with other volumes of Animal Cytogenetics; Arai, 2011; Blackmon et al., 2019; Gokhman, 2009; Graph odatsky et al., 2012; O’Brien et al., 2006; Olmo & Signorino, 2012; Perkins et al., 2019; Sylvester & Blackmon, 2020; Tree of Sex Consortium, 2014). Most of these sources were retrieved using Google Scholar and Google searches (i.e. animal AND chromosome number) conducted between 2019 and 2020. We used the GBIF taxonomy backbone, accessed using the rgbif R package version 3.5.2 (Chamberlain et al., 2020), to curate the ACC (Data S1). Three of the authors manually curated the full database. When available, the ACC list polymorphic counts per species. Analyses presented below are based on median haploid chromosome numbers per species.

2.2 | Examining the distribution of chromosome counts across the animal phylogeny

We first examined whether species in the ACC are distributed randomly, uniformly or clustered across the animal phylogeny. For this, we retrieved the species-level time-calibrated phylogeny for animals extracted from the Time Tree of Life (49,834 tips; Hedges et al., 2006). The original species-level phylogeny (49,834 tips) was then pruned down to include only species in the ACC (4,398 overlapping species). Next, we estimated the observed total sum of pairwise distances between species in the 4,398-tips tree, with species from the ACC. We then used the 49,834-tips tree to simulate a null distribution of the total branch lengths for trees of the same size as our data set (4,398 species). For this, we estimated the sum of pairwise distances between all pairs of species in 100 randomly generated subtrees each with 4,398 tips extracted from the 49,834-tips tree. Finally, we compared the observed sum of pairwise distances with the simulated distribution sum of pairwise distances. The relevant R script is provided in Data S2. We expected an observed value lower than the simulated distribution if chromosome numbers were clustered in our database. The opposite would be true if species were over-dispersed or uniformly distributed across the animal phylogeny. Pairwise distances were estimated using the cophenet.phylo function in the ape R package version 5.4-1 (Paradis & Schliep, 2019). Simulations were conducted in R using the base package 4.1.0 (R Core Team, 2021).

2.3 | Large-scale patterns of chromosome evolution in animals

We used three nonphylogenetic indices to summarize the patterns of chromosome evolution in animals and plants (Otto & Whitton, 2000). First, we calculated the incidence of polyploidy in animals and plants using the distribution of haploid chromosome numbers across species. This measure, also known as the polyploidy index (PI), summarizes the frequency of recent changes across species in haploid chromosome number that have occurred via polyploidization. Note that frequent polyploidization is expected to generate a large excess of even over odd haploid numbers: (#evens - #odds)/#evens (Otto & Whitton, 2000). We also determined the significance of the polyploidy index using a binomial test (sensu Otto & Whitton, 2000). Finally, we estimated the index of support (IS) on the PI to summarize the fraction of the analysed data set that must be independent for the PI to remain significantly different from zero. The lower the IS, the stronger the support for a nonzero PI. Second, we estimated the fraction of speciation events that are potentially associated with changes in chromosome numbers by totalling the minimum number of chromosome changes found within each genus and dividing this by the total number of speciation events within the same genus. Third, we estimated the frequency of speciation events involving polyploidization as the product between the PI and the fraction of speciation events associated with a change in chromosome number. We note that the assumptions made by these nonphylogenetic methods for estimating patterns of chromosome evolution are still to be tested (see Otto & Whitton, 2000). We also acknowledge that more widely used approaches for estimating parameters associated with changes in chromosome number between species are currently available (e.g. ChromEvol: Glick & Mayrose, 2014; Mayrose et al., 2010; ChromoSSE: Freyman & Höhna, 2018). However, these approaches are explicitly based on phylogenetic information, which is not available for most of the species in the database.

We compared between animals and plants (ferns and angiosperms: Rice et al., 2015) the estimated values of the PI, the fraction
of speciation events that are potentially associated with changes in chromosome numbers and the frequency of speciation events involving polyploidization. For this, we estimated the 95% confidence interval (based on percentiles) for each of the three indexes outlined above for animals and plants using a total of 100 bootstrap replicates. The relevant R script is provided in Data S3.

3 | RESULTS

3.1 | The ACC

We present the first database of chromosome counts for the entire animal kingdom (Table 1). The current version of the ACC includes chromosome counts for 14,524 species across 5,982 genera, 1,279 families, 263 orders, 61 classes and 21 phyla (Figure 1; Table 1). Among the 21 phyla sampled in the current release of the ACC, only seven groups have chromosome counts for ≥1% of their extant diversity: Phoronida (12.5% of total species richness), Chordata (6.5%), Chaetognatha (2.4%), Nematomorpha (2.1%), Sipuncula (1.5%), Entoprocta (1.2%) and Platyhelminthes (1%). Species in the database have between 1 and 22 haploid counts with an average of 1.14 counts. Chromosome numbers in the current release of the ACC range from \( n = 1-191 \), with a median chromosome count of 13, a mean of 15.36 and a mode of 12.

3.2 | Examining the distribution of chromosome counts across the animal phylogeny

Our results indicate that the chromosome counts in the ACC database are clustered to specific branches of the animal Tree of Life. The sum of pairwise distances between species in our database \( (2.008\times10^{10}) \) was significantly smaller than the confidence interval for the null distribution of total pairwise distances between species in phylogenetic trees of the same size to our data set (95% CI, \( 2.297 \times 10^{10}-2.332 \times 10^{10} \)). Therefore, chromosome counts in the ACC are not randomly distributed across the animal phylogeny.

3.3 | Large-scale patterns of chromosome evolution in animals

We compared the evolution of chromosome counts between animals and plants using a nonphylogenetic approach. Different evolutionary processes are responsible for producing similar distributions

<table>
<thead>
<tr>
<th>Phylum</th>
<th>Species</th>
<th>Number of ( n ) counts</th>
<th>Unique ( n ) counts</th>
<th>Mean ( n )</th>
<th>Median ( n )</th>
<th>Mode</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acanthocephala</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>3–8</td>
</tr>
<tr>
<td>Annelida</td>
<td>72</td>
<td>88</td>
<td>28</td>
<td>15</td>
<td>13</td>
<td>16.18</td>
<td>3–110</td>
</tr>
<tr>
<td>Arthropoda</td>
<td>9,755</td>
<td>12,369</td>
<td>85</td>
<td>13</td>
<td>11</td>
<td>10</td>
<td>1–191</td>
</tr>
<tr>
<td>Bryozoa</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>6.5</td>
<td>11.673</td>
<td>3–11</td>
</tr>
<tr>
<td>Chaetognatha</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>8.8</td>
<td>9</td>
<td>9</td>
<td>8–9</td>
</tr>
<tr>
<td>Chordata</td>
<td>4,134</td>
<td>4,885</td>
<td>71</td>
<td>21</td>
<td>20</td>
<td>18</td>
<td>1–125</td>
</tr>
<tr>
<td>Cnidaria</td>
<td>18</td>
<td>35</td>
<td>14</td>
<td>11</td>
<td>11</td>
<td>6.12</td>
<td>5–31</td>
</tr>
<tr>
<td>Dicyemida</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>–</td>
</tr>
<tr>
<td>Echinodermata</td>
<td>26</td>
<td>55</td>
<td>23</td>
<td>15</td>
<td>17</td>
<td>18</td>
<td>4–32</td>
</tr>
<tr>
<td>Entoprocta</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>9.5</td>
<td>9.5</td>
<td>11.8</td>
<td>8–11</td>
</tr>
<tr>
<td>Mollusca</td>
<td>122</td>
<td>195</td>
<td>27</td>
<td>18</td>
<td>17</td>
<td>29</td>
<td>5–31</td>
</tr>
<tr>
<td>Nematoda</td>
<td>35</td>
<td>88</td>
<td>18</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>1–24</td>
</tr>
<tr>
<td>Nematomorpha</td>
<td>7</td>
<td>12</td>
<td>6</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1–8</td>
</tr>
<tr>
<td>Nemertea</td>
<td>10</td>
<td>22</td>
<td>16</td>
<td>14</td>
<td>17</td>
<td>16</td>
<td>1–25</td>
</tr>
<tr>
<td>Phoronida</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>9</td>
<td>9</td>
<td>12.6</td>
<td>6–12</td>
</tr>
<tr>
<td>Platyhelminthes</td>
<td>306</td>
<td>403</td>
<td>21</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>2–26</td>
</tr>
<tr>
<td>Porifera</td>
<td>8</td>
<td>16</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>8</td>
<td>4–23</td>
</tr>
<tr>
<td>Rotifera</td>
<td>4</td>
<td>16</td>
<td>13</td>
<td>15</td>
<td>13</td>
<td>13,12,24</td>
<td>6–26</td>
</tr>
<tr>
<td>Sipuncula</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>–</td>
</tr>
<tr>
<td>Tardigrada</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>–</td>
</tr>
<tr>
<td>Xenacoelomorpha</td>
<td>7</td>
<td>22</td>
<td>13</td>
<td>13</td>
<td>14</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>

Note: We indicate the number of species in the database, total and unique number of haploid counts, along with the mean, median, mode, and range of haploid chromosome numbers.
of chromosome numbers in animals and flowering plants (Figure 1). For instance, we estimated that 11% of the changes in chromosome number among animals occurred via polyploidization (median $PI = 11.315\%$, 95% CI = 8.354–13.890\%, $p < .001$, IS = 0.3\%; Table 2) and 29% of the changes in chromosome number occurred with speciation events (95% CI, 28.274%–30.183\%). Relative to plants, changes in chromosome number via polyploidization were more than two times less frequent in animals compared with either angiosperms or ferns (angiosperms: median $PI = 29.45\%$, 95% CI = 28.182–31.051\%, $p < .001$, IS = 0.1\%; ferns: median $PI = 46.245\%$, 95% CI = 43.257–49.598\%, $p < .001$, IS = 3.9\%). Similarly, the percentage of speciation events associated with polyploidy was more than three times higher in plants than in animals (animals: median = 3.29\%, 95% CI = 2.419–3.995\%; angiosperms: median = 9.85\%, 95% CI = 6.874–7.610; ferns: median = 14.465\%, 95% CI = 13.274–15.907\%). However, the frequency of speciation events associated with changes in chromosome number in animals (median = 29.120\%, 95% CI = 28.274–30.183) was higher than in angiosperms (median = 24.530\%, 95% CI = 23.959%–25.035\%) but similar to that in ferns (median = 31\%, 95% CI = 29.204–33.369\%). Overall, in addition to highlighting the complexity of chromosome evolution across different lineages, these indexes suggest that different processes are responsible for the similarity in the distribution of chromosome counts between animals and flowering plants (Figure 1; Table 2).

4 | DISCUSSION

4.1 | Examining the distribution of chromosome counts across the animal phylogeny

We present the most extensive collection of chromosome counts for animals currently available. Our database contains counts for 14,524 species, representing ~1% of all described animal species. Although this is the largest and most comprehensive database yet available for animal chromosome counts, many phyla are not represented at all. We still do not know the number of chromosomes—the most basic description of their genomes—for entire clades of animals. Specifically, the ACC still lacks any chromosome counts for 11 of the 27 commonly recognized animal phyla: Acoela, Brachiopoda, Bryozoa, Ctenophora, Gastrotricha, Gnathostomulida, Hemichordata, Kinorhyncha, Onychophora, Placozoa and Xenoturbellida.

Overall, the ACC highlights not only the lack of population-level counts for many animal species but also the extensive lack of basic information on chromosome numbers for even the entire phyla. The taxonomic sampling in the ACC stands in contrast to the information available on plant chromosomes. The plant-focussed Chromosome Count Database has more than 300,000 chromosome counts for nearly 200,000 species of green plants (Rice et al., 2015). However, unlike in animals, plant cytological research was driven in the early 20th century by excitement over polyploidy (Barker et al., 2016), and the result is an extensive catalog of information on plant chromosomes. It is clear from our compilation of data that there is a need for basic cytological research across the animal tree of life. Future

![Distribution of haploid chromosome numbers for animals, angiosperms and ferns](image)

**TABLE 2** Summary statistics for haploid chromosome numbers compiled for ferns, flowering plants, and animals

<table>
<thead>
<tr>
<th>Lineage</th>
<th># Species</th>
<th># Even</th>
<th>Mode (%)</th>
<th>PI % (95% CI)</th>
<th>$p$-value [IS%]</th>
<th>SpeChangeChrom % (95% CI)</th>
<th>SpePolyp % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animals</td>
<td>14,524</td>
<td>9,654</td>
<td>0.01</td>
<td>11.315 (8.354–13.890)</td>
<td>4E–324 (0.3)</td>
<td>29.120 (28.274–30.183)</td>
<td>3.29 (2.419–3.995)</td>
</tr>
<tr>
<td>Ferns</td>
<td>1,729</td>
<td>1,092</td>
<td>14.2</td>
<td>46.245 (43.257–49.598)</td>
<td>1E–27 (3.9)</td>
<td>31.385 (29.204–33.369)</td>
<td>14.465 (13.274–15.907)</td>
</tr>
</tbody>
</table>

Note: We summarize the number of species in each database under the # Species column. The number of even haploid chromosome counts is indicated under the # Even column for each lineage. The per cent of species in the database with the modal haploid chromosome count is indicated under the ModePercent column. The polyplody index (PI) column, along with its associated $p$-value, and index of support (IS) column are indicated in the table for each lineage. Finally, we summarize for each clade both the fraction of speciation events associated with chromosome changes (SpeChangeChrom column) and polyploidization (SpePolyp column). We follow Otto and Whitton (2000) for PI, SpeChangeChrom and SpePolyp equations and estimates for plants. We present the median and confidence interval for each of the three indexes based on 100 bootstrap replicates.
releases of the ACC should aim to expand the taxonomic sampling within the already sampled phyla, but also focus on compiling counts for the currently unsampled ones.

**4.2 | Why is the range of chromosome numbers similar between animals and flowering plants?**

We observed a similar distribution in chromosome numbers between animals and flowering plants. Although this could be due to chance, the similar distribution of numbers suggests that there may be selection for \( n = 9–12 \) chromosomes in both of these lineages. Ferns, in contrast, had much higher numbers of chromosomes with a peak at \( n = 41 \). Chromosome number evolution in the ferns is well recognized as exceptional relative to other eukaryotes (Barker & Wolf, 2010; Otto & Whitton, 2000; Wood et al., 2009), but the similarity in the distribution of chromosome numbers between flowering plants and animals was not previously recognized. We are not aware of previous studies highlighting the similarity in the distribution of chromosome counts between animals and flowering plants. However, our study does indicate that the drivers of the similar distribution of chromosome counts in animals and plants are not equivalent. Future studies should examine the factors that underlie this remarkable and unexpected similarity in the distribution of chromosome counts between these two groups.

**4.3 | Why are changes in chromosome number related to speciation and not polyploidy in animals?**

Although animals and flowering plants have similar distributions of chromosome numbers, we found that the underlying processes were different. In particular, polyploidy is more strongly associated with changes in chromosome number and speciation in plants than in animals. This is a well-known difference in plant and animal speciation (Coyne et al., 1993; Gregory & Mable, 2005; Muller, 1925; Otto & Whitton, 2000; Sites & Moritz, 1987; Stebbins, 1958), and our results confirm this long-standing observation. The specific mechanisms explaining why polyploidy is less common in animals probably relate to differences in sex determination (Muller, 1925), meiotic disjunction mechanisms (Macgregor, 1993) and embryology (Stebbins, 1950; von Wettstein, 1927), or even the frequency of self-fertilization (White, 1973) and the absence of degenerate sex chromosomes (Orr, 1990). Ultimately, the relative rarity of polyploidy among animals is still an open question (Gregory, 2011), and expanding this database will provide the opportunity for further comparative analyses to test these hypotheses. In contrast, the frequency of speciation-related changes in chromosome number was similar between animals and flowering plants, suggesting that changes in karyotype and chromosome number may have similar impacts on fertility in both lineages.

**5 | CONCLUSIONS**

Here, we present the largest data set of chromosome counts for animals. The ACC summarizes haploid counts for ~15,000 species, representing ~1% of the extant animal diversity. Note that our results suggest that the distribution of haploid counts in animals is robust to the inclusion of new observations. We found a remarkable similarity in the distribution of chromosome counts between animals and plants. Overall, the similar distribution of chromosome numbers between animals and flowering plants may be best explained by selection driving numbers towards a narrow range. Nonphylogenetic analyses suggest that the similarity in the distribution of chromosome counts between animals and plants is driven by a different process (e.g. polyploidy is more often related to speciation in plants than in animals). Future studies should examine the factors influencing the distribution of chromosome counts between animals and plants. Additionally, extensive data collection of chromosome numbers is still a priority for animals, where more than 99% of species are yet to be sampled and where even counts for entire phyla are still lacking in the database.

**ACKNOWLEDGMENTS**

Because the ACC is extensively founded on existing chromosome counts compiled by multiple research groups or summarizing counts as a by-product, we thank the efforts of hundreds of researchers for making data available for others. This database is intended to serve as a quick and highly accessible reference for further research.

**CONFLICT OF INTEREST**

The authors have no conflict of interest to declare.

**AUTHOR CONTRIBUTIONS**

CRP, CAM and SHZ collected and curated data; CRP wrote the manuscript. CRP, MSB, CAM and MSB revised and approved the final version of the manuscript.

**DATA AVAILABILITY STATEMENT**

All the counts are available at https://cromapa94.github.io/ACC/. Updates to the database will be part of two of the authors’ (C.R.P. and M.S.B.) long-term research goals. Contributions (new or even previously collected chromosome counts) can be directly submitted by following the instructions on the website. We only require submissions to provide at least the following three details: (i) species name, (ii) chromosome count(s) and (iii) a given reference. Information can be accessed through the corresponding website, and counts can be downloaded as a CSV. The current release of the ACC is also available through Zenodo (10.5281/zenodo.4635457). Updates to the database will be performed yearly after the publication of this study.

**PEER REVIEW**

The peer review history for this article is available at https://pubon/10.1111/jeb.13884.


**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Román-Palacios, C., Medina, C. A., Zhan, S. H., & Barker, M. S. (2021). Animal chromosome counts reveal a similar range of chromosome numbers but with less polyploidy in animals compared to flowering plants. *Journal of Evolutionary Biology*, 00, 1–7. https://doi.org/10.1111/jeb.13884
Animal chromosome counts reveal a similar range of chromosome numbers but with less polyploidy in animals compared to flowering plants

Cristian Román-Palacios | Cesar A. Medina | Shing H. Zhan | Michael S. Barker

1Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA
2Department of Neuroscience, University of Arizona, Tucson, AZ, USA
3Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC, Canada

Correspondence
Cristian Román-Palacios, Center for Biodiversity, Temple University, Philadelphia, PA 19122, USA.
Email: cromanpa94@gmail.com

Graphical Abstract
The contents of this page will be used as part of the graphical abstract of html only. It will not be published as part of main article.

We describe a previously undocumented biological pattern—animals and flowering plants exhibit remarkably similar distributions of haploid chromosome numbers.