

APPLICATION

phenotools: An R package for visualizing and analysing phenomic datasets

Chad M. Eliason^{1,2}  | Scott V. Edwards³  | Julia A. Clarke¹ 

¹Department of Geological Sciences, University of Texas Austin, Austin, Texas

²Grainger Bioinformatics Center, Field Museum of Natural History, Chicago, Illinois

³Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts

Correspondence

Chad M. Eliason

Email: celiason@fieldmuseum.org

Julia A. Clarke

Email: julia_clarke@jsg.utexas.edu

Funding information

Gordon and Betty Moore Foundation, Grant/Award Number: 4489; Division of Earth Sciences, Grant/Award Number: NSF EAR 1355292

Handling Editor: Natalie Cooper

Abstract

1. Phenotypic data are crucial for understanding genotype–phenotype relationships, assessing the tree of life and revealing trends in trait diversity over time. Large-scale description of whole organisms for quantitative analyses (phenomics) presents several challenges, and technological advances in the collection of genomic data outpace those for phenomic data. Reasons for this disparity include the time-consuming and expensive nature of collecting discrete phenotypic data and mining previously published data on a given species (both often requiring anatomical expertise across taxa), and computational challenges involved with analysing high-dimensional datasets.
2. One approach to building approximations of organismal phenomes is to combine published datasets of discrete characters assembled for phylogenetic analyses into a phenomic dataset. Despite a wealth of legacy datasets in the literature for many groups, relatively few methods exist for automating the assembly, analysis, and visualization of phenomic datasets in phylogenetic contexts. Here, we introduce a new R package `phenotools` for integrating (fusing original or legacy datasets), curating (finding and removing duplicates) and visualizing phenomic datasets.
3. We demonstrate the utility of the proposed toolkit with a morphological dataset for flightless birds and two morphological datasets for theropod dinosaurs and provide recommendations for character construction to maximize accessibility in future workflows. Visualization tools allow rapid identification of anatomical sub-regions with difficult or problematic histories of homology.
4. We anticipate these tools aiding automation of the assembly and visualization of phenomic datasets to inform evolutionary relationships and rates of phenotypic evolution.

KEYWORDS

computational biology, homology, machine learning, morphological traits, phenomics, phylogenetics

1 | INTRODUCTION

Discretized anatomical, behavioural, or life-history traits previously used in phylogenetics, an important form of phenomic data, may be crucial for understanding the genetic bases of complex traits (Lee

& Palci, 2015), uncovering relationships among extinct and extant organisms (Gavryushkina et al., 2015), placing fossils for dating molecular phylogenies (Jarvis et al., 2014), and elucidating changes in morphological disparity through time (Clarke & Middleton, 2008). Compared to genomic data, large-scale phenotyping of organisms

presents several challenges. First, phenome assembly and development of analytical tools for phenomic data lag behind those for genomic data (Deans et al., 2015). Second, phenomic data are difficult to collect, requiring both anatomical and taxonomic expertise in assessing homology (Patterson, 1982) and coding variability in traits (Lee & Palci, 2015). This problem is exacerbated by the dwindling number of expert taxonomists working on anatomical descriptions of extant and extinct groups (Kemp, 2015). One way forward is to utilize existing phenomic data in the form of published descriptions, trait databases or natural history collections (Deans et al., 2015). However, a third major challenge in phenomics is converting these data sources into computable phenotypes (Burleigh et al., 2013; Deans et al., 2015). Assessing character quality on a case-by-case basis is a rate-limiting step in phenomics, especially for large (>1,000 character) datasets (e.g. O'Leary & Kaufman, 2011).

Phenomic datasets can be constructed using either manual or automated approaches. Manual approaches involve utilizing trait ontologies (e.g. Dahdul et al., 2012) to assign metadata to narrative character descriptions, possibly using Phenex (Balhoff et al., 2014). Individual phenomes are then concatenated into a phenomic dataset (set of taxa and characters), for example using MorphoBank (O'Leary & Kaufman, 2011). Automated approaches use natural language processing to extract characters from text-based taxonomic descriptions (Cui, 2012; Endara, Cui, & Burleigh, 2018). Existing approaches output 'raw' character matrices (Cui, 2012), and it is up to the user to identify character states and evaluate the utility and homology of proposed characters and character states. Given the wealth of phenomic data available in the literature (Deans et al., 2015), there is a critical need for tools aiding in assembling and evaluating phenomic datasets.

Here, we present a pipeline to reproducibly assemble and visualize phenomic datasets with the goals of making character annotation faster (through automation) and homology assessment easier (through network representation of character description similarity). This pipeline is implemented in a new R package `phenotools` that inputs published datasets and outputs a merged phenomic dataset. We demonstrate the utility of this pipeline using a newly assembled phenomic dataset for palaeognathous birds (Aves: Palaeognathae), as well as two existing morphological datasets for theropod dinosaurs. Finally, we provide recommendations for character construction and future directions for phenome assembly.

2 | OVERVIEW OF THE PHENOTOOLS PACKAGE

`phenotools` provides functionality for inputting phenomic datasets; merging datasets into a single phenomic dataset based on a common set of taxa; annotating characters based on fuzzy text searching/trait ontologies; filtering overlapping phenomic characters and taxa (e.g. merging taxa into supraspecific terminals); visualizing similarities in character description; and outputting phenomic datasets to common phylogenetic data formats for downstream analysis (Figure 1).

2.1 | Importing phenomic datasets

Users can input data into `phenotools` using `read.nex` in one of the three ways, depending on the input file format. First, individual NEXUS files (Maddison & Maddison, 2018) are input as is. Second, separate data matrices and character lists (as text files) can be read in. Third, data matrices can be extracted from within PDF files, although this feature is experimental and not fully implemented. Users may prefer to cut and paste character matrices into other programs (e.g. Mesquite), but this can result in user-generated errors and, we believe, not ideal as part of a reproducible, streamlined workflow.

2.2 | Extracting ontologies from character descriptions

Anatomical traits, like most phenotypic traits, are typically hierarchical and comprised of distinct subunit parts (McShea & Venit, 2001). However, character statements are not always written to express this hierarchy (Figure 2). Our automated approach implemented in `generate_ontology` constructs a trait ontology (i.e. a controlled vocabulary for describing phenotypes) from narrative character descriptions with hierarchical structure defined by comma-separated terms (Figure 2b). We have also implemented a manual approach in `read_ontology` that inputs a user-defined ontology as a text file. To link characters to terms in the ontology, we first simplify terms using the Schinke algorithm (Schinke, Greengrass, Robertson, & Willett, 1996), useful for Latin terms common in anatomical datasets (e.g. 'humerus' becomes 'humer'). Metadata is then associated with the characters based on the strongest match (i.e. longest set of overlapping characters or terms) between an ontological term and a given character (Figure 2b) using `traitlink`.

2.3 | Finding overlapping characters

Different authors have different ways of writing characters. For characters used by subsequent authors that are reproduced verbatim, finding duplicates is straightforward (Ramírez et al., 2007). In cases where characters are written in different languages, the process of locating duplicated characters is challenging. We dealt with this by using roots in Latinate languages. For languages other than English or Latinate ones, identifying duplicate character descriptions will still be an issue. Similar characters may also be alternative descriptions of the same morphology. For example, one author may describe a structure as present or absent, while another author may describe more nuanced variation in shape. Authors may also differ in the taxonomic sample, or in character states and scorings used to describe a trait.

To begin to address these issues, we implemented four approaches for identifying potential overlapping characters in `duplicated.nex`. The first fuzzy text approach (`opt='fuzzy'`) identifies potential matching characters based on similarity in characters and character states. Positional terms (e.g. 'dorsal'), white space, punctuation and comments are removed prior to

matching, and a range of text distance metrics can be specified (e.g. see van der Loo, 2014). For example, `method='jw'` uses Jaro-Winkler distances, or the percentage of matched characters between two strings. Parts of character descriptions can be further up-weighted. For example, locators (terms that point to a particular feature, such as an anatomical structure) are often specified early in character statements and delimited by commas (Figure 2b). This information can be leveraged by specifying `weighting=c(2,1,1)`. The second term-based approach (`opt='terms'`) searches the entire dataset to determine word frequencies. Common terms are then down-weighted relative to less-common, anatomical terms (e.g. 'humerus'). Various community detection algorithms (Csardi & Nepusz, 2006) can be specified with the `cluster` argument to identify groups of similarly worded

character statements (Figure 2a). The third text-searching approach (`opt='comments'`) tests if character statements include comments referencing other characters. For example, Figure 2a shows three character statements comprising a single cluster of duplicated characters, with the original character described by Lee, Feinstein, and Cracraft (1997) subsequently modified and used in phylogenetic analyses by Bourdon, Ricqles, and Cubo (2009) and Worthy and Scofield (2012). The fourth distance-based approach (`opt='traitcor'`) determines pairwise correlations between characters based on Hamming distances (number of character state differences between two characters) for binary characters (`method='hamming'`) or polychoric correlations (strength of association between two ordinal variables) for multistate characters (`method='polychor'`; see Supplementary Methods for details).

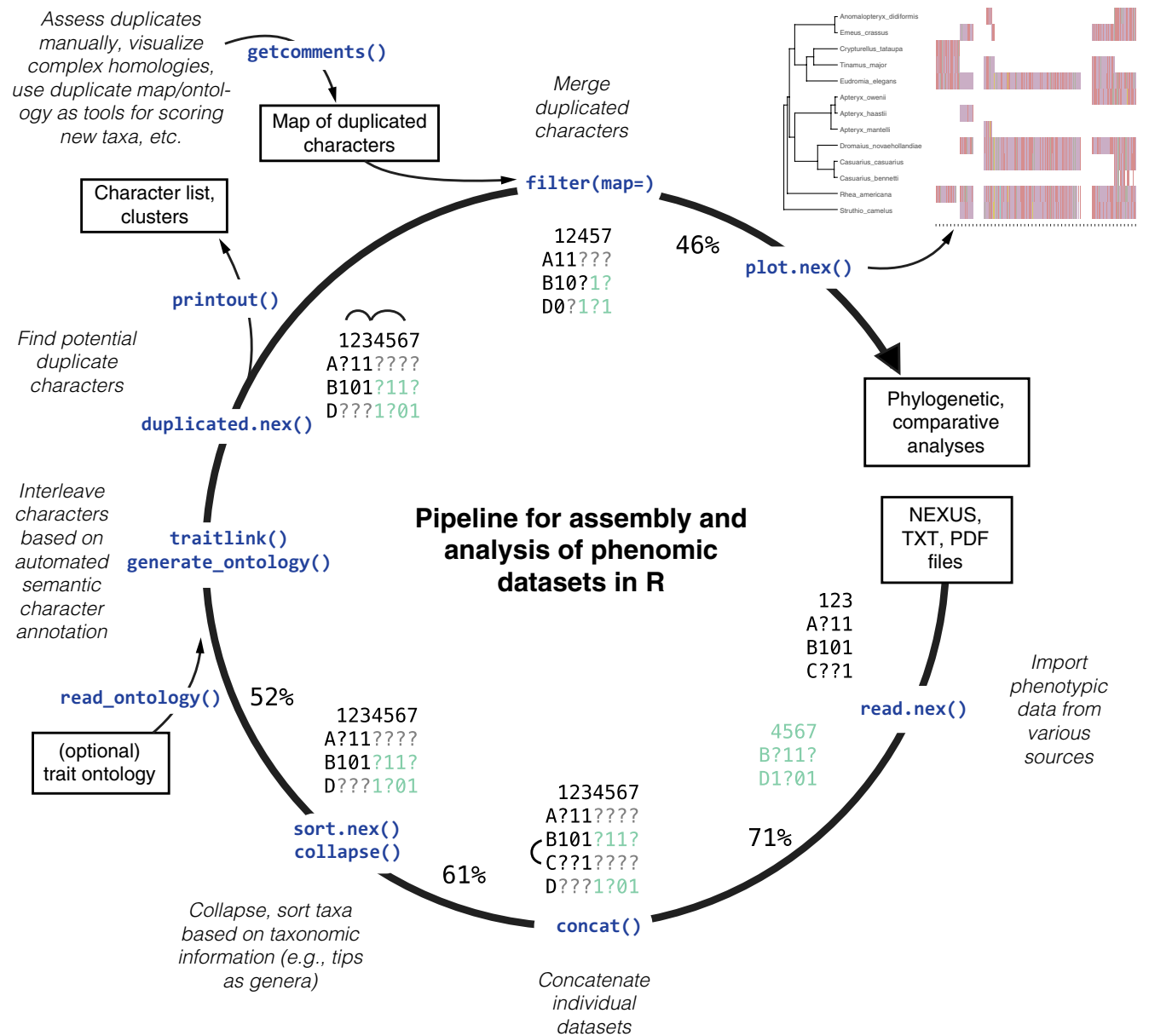


FIGURE 1 Pipeline for phenomic dataset assembly in the `phenotools` R package. R functions indicated along inner circle in dark blue text. Spin-offs indicate actions that can be performed on the dataset at a given point in the pipeline. Numbers around circle indicate % missing data

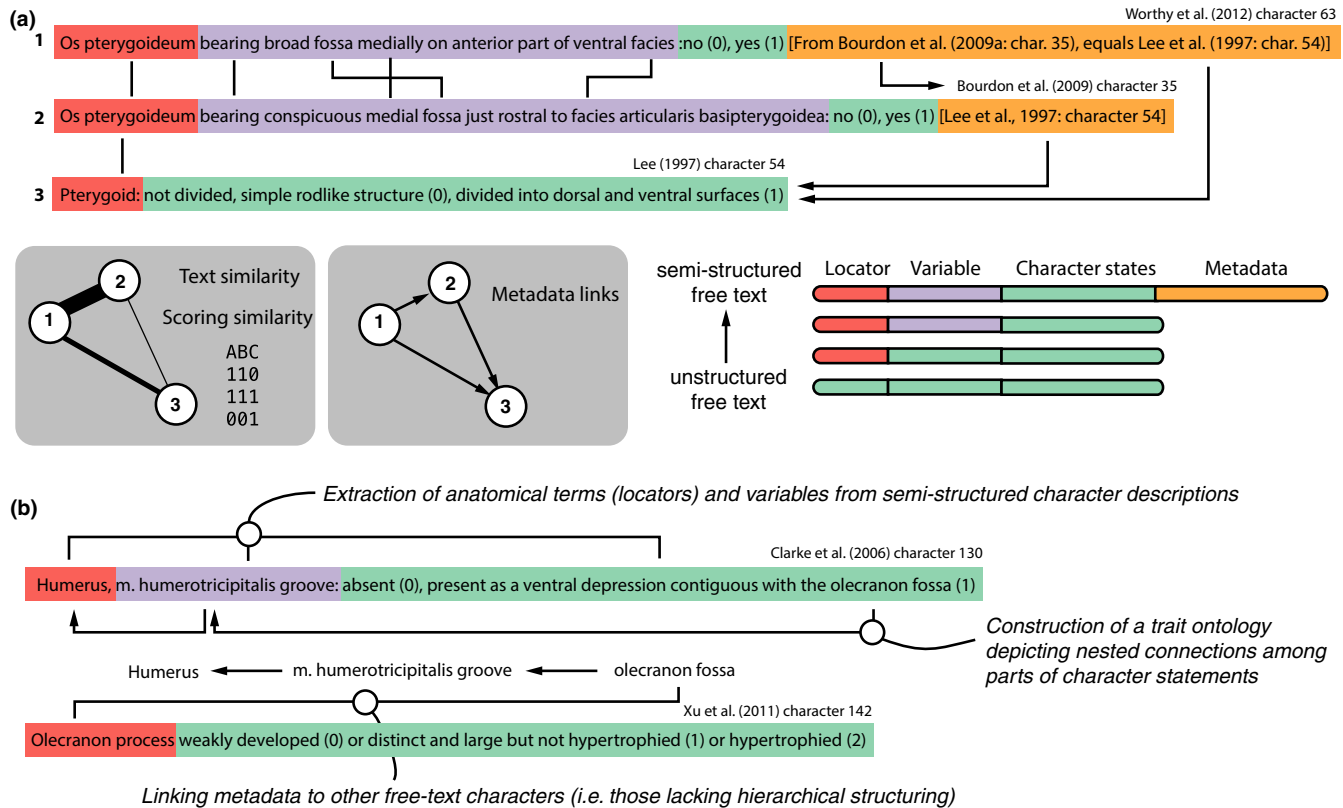


FIGURE 2 Character similarity networks and character annotation. (a) Anatomy of a character statement. Arrows indicate two approaches for comparing characters: fuzzy text matching of terms (lines without arrowheads) and text searching within characters (lines with arrowheads). Network diagrams in gray boxes show connections among characters using different approaches, with thicker lines indicating more similar characters. (b) Automated semantic character annotation with `generate_ontology`. Character statements can be structured along a continuum, with semi-structured characters (Clarke, Zhou, & Zhang, 2006) presenting greater opportunity for automated phenomic dataset assembly

Character correlations are useful because, in addition to similarity in character statement terminology, overlapping characters might be expected to show similar scores for a set of taxa. Strong correlations could therefore indicate either logical dependence or biologically relevant, functional or developmental dependence among characters (e.g. phenotypic integration). Assessing and interpreting these alternative scenarios for a set of similar characters would be up to the user.

3 | WORKED EXAMPLE: A PHENOMIC DATASET FOR PALAEOGNATHOUS BIRDS

After converting and/or downloading published datasets (see Supplementary Methods) as NEXUS files, we loaded them into R using:

```
files <- list.files("datasets/palaeognaths",
  pattern = "nex")
nexlist <- lapply(files, read.nex)
```

We then concatenated datasets with:

```
allnex <- concat(nexlist, taxa = taxonlist)
```

Since some taxa are subspecies, we chose to use `collapse` to merge subspecies (e.g. *Rhynchotus rufescens pallescens*) and rename

some species (e.g. *Pterocnemia pennata*; see R script available at Dryad for details):

```
final <- collapse(allnex, map = list("Pterocnemia
pennata" = "Rhea pennata", "Rhynchotus rufescens
pallescens" = "Rhynchotus rufescens", method =
"merge"))
```

Finally, we sorted by taxa using:

```
final.sorted <- sort.nex(final, by="taxlabels")
```

The resulting dataset contained 4,333 characters and 122 taxa. After concatenating datasets, relabeling taxa and removing invariant characters with `remove.invar`, the dataset contained 2,506 characters for 122 taxa, with 87% missing data. This process took ~10 s on 2 GHz Intel Core i7 MacBook Pro laptop with 8 GB of RAM.

At this point in the pipeline, we anticipate users will want to export clusters of similar character descriptions and begin to address character homology issues using:

```
final.dups <- duplicated.nex(final.sorted,
  opt = "terms")
```

Clusters of similar character descriptions can be visualized with shared terms among characters colour-coded using the `printout` function (Figure 3b, S1):

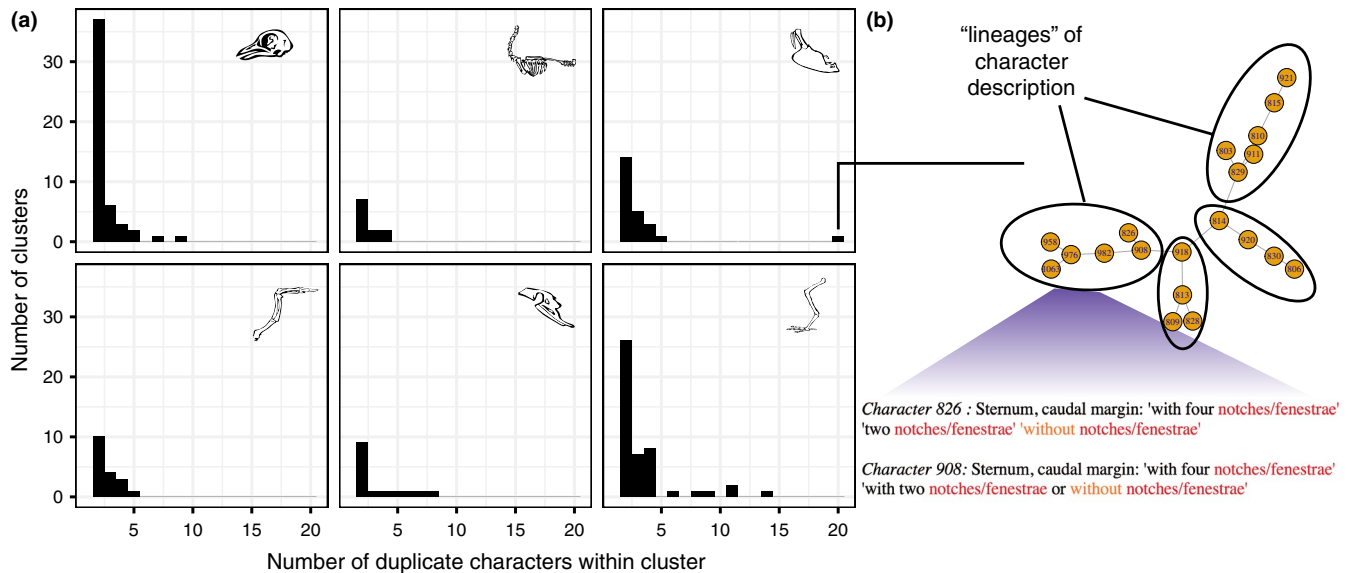


FIGURE 3 Visualizing character overlap. (a) Number of clusters with similar character descriptions (based on manual assessment of characters), grouped by anatomical region for the palaeognath dataset. (b) Visualization of a highly connected group of characters with distinct ‘lineages’ of similarly worded characters. A cluster of characters output with the `printout` function shows colour-coded terms that place groups of similar characters in a given cluster

```
printout(final.dups, file="clusters.html")
```

After manually assessing this output and adding comments (e.g. ‘DD 100’ to indicate that a focal character is a duplicate of character 100), `getcomments` can import these annotations into R or, alternatively, duplicates can be input as a spreadsheet with columns for focal and duplicated characters. Characters can be automatically merged or dropped (see Supplementary Methods), resulting in a clean dataset for downstream phylogenetic or comparative analyses:

```
final.clean <- filter(final.sorted, map="dups.csv")
```

Finally, we can plot (e.g. alongside a phylogeny; Figure S2):

```
plot.nex(x=final.sorted, phy=phy)
```

or export the cleaned dataset as a NEXUS file:

```
write.nex(final.clean, file="supermatrix.nex",
          format="nexus")
```

4 | BENCHMARK

To assess the performance of our method of annotating characters, we determined the proportion of characters correctly linked to a given body region. For the theropod datasets, 71% of the characters in Xu, You, Du, and Han (2011) included words in a trait ontology for birds (Baumel & Witmer, 1993), compared to 81% of the characters from Brusatte, Lloyd, Wang, and Norell (2014) (Table 1). Between 57% and 73% of characters were assigned to the correct body region across datasets (Table 1). The two theropod datasets differed in character annotation accuracy (Table 1), likely because of the richer character descriptions in Brusatte et al. (2014), and thus greater potential for locating important anatomical terms in the trait ontology.

Several characters from Xu et al. (2011) are currently written with little structured information about the precise location of described variability. The improved rate of character annotation in the palaeognath dataset (Table 1) likely stems from differences in the set of anatomical terms used to describe crown birds versus theropod dinosaurs. To assess the performance of our four methods for finding overlapping characters, we manually verified whether proposed sets of characters were true or false duplicates. The text-searching method had the highest precision, while the fuzzy text-matching method had the highest recall (Table 2). The term-based approach had moderate performance at both precision and recall and outperformed the fuzzy text-matching approach in its ability to locate duplicates (Table 2). Importantly, the cut-off for these algorithms can be adjusted by adjusting the cut-off argument to optimize precision and recall. See Supplementary Methods for details.

5 | DISCUSSION

Automated processes for phenomic dataset assembly will facilitate progress towards phenomes for all organisms, although human expertise will be needed in error checking and homology evaluation. `phenotools` provides a set of tools and a pipeline for this process (Figure 1) that work alongside existing tools. For example, a user might form a phenomic dataset from published datasets, generate an ontology, remove potential duplicates and then upload this final dataset to MorphoBank (O’Leary & Kaufman, 2011) to make the data accessible to other researchers. Programs currently exist to accomplish various subtasks in phenomics (Balhoff et al., 2014; Liu, Endara, & Burleigh, 2015; Maddison & Maddison, 2018; O’Leary & Kaufman, 2011). A core novelty of our approach is a reproducible pipeline for phenomic dataset

TABLE 1 Semantic character annotation performance

Dataset	Total characters	Linked to avian trait ontology	Correctly annotated (%) ^a	Cluster size, mean (range) ^b	Words per character description, mean (range)
Palaeognathae (this study)	715	652 (91%)	476 (73%)	3.6 (2–20)	10 (2–45)
Brusatte et al. (2014)	853	669 (81%)	412 (67%)	4.3 (2–22)	12 (1–83)
Xu et al. (2011)	374	265 (71%)	151 (57%)	3.0 (2–8)	8 (1–34)

Note: Formatted trait ontology of Baumel and Witmer (1993) available at github.com/celiason/phenotools. Character descriptions and character states combined in Brusatte et al. (2014) and Xu et al. (2011), whereas these are separate for the palaeognath dataset. Results are shown for character descriptions only. The palaeognath dataset was filtered to include only binary characters and completely scored taxa.

^aPercent correctly annotated characters indicate the proportion of characters correctly linked to a given body region. For example, a character of the cranium being located with other cranial characters.

^bCluster size refers to the number of characters in a group of similar character descriptions.

TABLE 2 Duplicate finding performance

Method	False neg.	True pos.	False pos.	Precision TP/(TP + FP)	Recall TP/(TP + FN)
Fuzzy text-matching	318	39	689	5.4%	10.9%
Weighted term-based search	295 ^a	62 ^b	206 ^c	23.1%	17.4%
Text-searching	337	20	8	71.4%	5.6%
Distance-based: Hamming distance	246	111	8,743	1.3%	31.1%
Distance-based: Polychoric distance	300	57	4,614	1.2%	16.0%

Note: Calculated precision and recall for the palaeognath hindlimb dataset (233 characters, 122 taxa, 86.5% missing data). Precision indicates the proportion of potential duplicates that are true duplicates (verified manually) and recall indicates the proportion of true duplicates flagged as potential duplicates.

^aExamples of false negatives: (1) 'Tibiotarsus, condylus lateralis triangular in shape and separated from condylus medialis by shallow, proximodistally narrow incisura intercondylaris: no, yes' and (2) 'Tibiotarsus, cranial aspect, relative proximodistal length condylus medialis to minimum proximodistal length incisura intercondylaris distal to its rim above the impressio ligamenti intercondylaris: incisura short, less than half length of condylus medialis, incisura long, much more than half of length of condylus medialis'.

^bExamples of true positives: (1) 'Tibiotarsus, condylus medialis with very deep ligamental pit: no, yes' and (2) 'Tibiotarsus, medial side of the internal condyle: slight depression near anterior margin, deep pit in anterior margin and a groove along the posterior margin'.

^cExamples of false positives: (1) 'Tarsometatarsus, number of hypotarsal sulci: three, two, one' and (2) 'Tarsometatarsus, number of subhypotarsal ridges: three or two, one'.

assembly (Figure 1). Automating the phenome assembly process minimizes error and allows future researchers to retrace the process from start (individual NEXUS files) to finish (phylogenetic analyses).

Our network-based approach facilitates rapid discovery of important terms that can be used to modify or build new datasets; visualizes clusters of overlapping characters to make it easier for researchers to identify overlapping characters and assess homology; and is flexible in its ability to incorporate different text algorithms for character similarity that will be applicable to diverse taxa. The rich HTML output with colour-coded terms produced by `printout` (Figure 3b) locates shared terms among characters, and could facilitate dataset quality control (e.g. identifying overlapping characters after concatenation of individual datasets) or aid in the construction of new datasets (e.g. by identifying a controlled vocabulary). Information about the temporal, logical or ontological relationships among characters, as obtained from manual assessment of datasets or using `phenotools`, can also reveal temporal trends in rates of character gains and losses over time (Figure S3) and spatial distribution of characters across related datasets (Figure 3, S5). These

tools are more of interest in assessing the trends in character use and phylogenetic practices, but comparative network analysis using the theropod dinosaur datasets (Figure S6) illustrates how these visualization tools can be used with existing phenomic datasets.

Going forward, we echo the suggestion of Deans et al. (2015) that characters be written to maximize computer readability. For example, characters written in a hierarchical fashion (see Figure 2) make it easier to automatically construct trait ontologies using `generate_ontology`. We also suggest that, where possible, authors avoid acronyms within character or state labels, as well as include comments in square brackets to facilitate identification of overlapping characters. Our weighted term-based approach captured only 23% of duplicates (Table 1), highlighting the ongoing need for taxonomic and anatomical expertise in phenomics. However, given manual assessment took ~20 min per character (Nick M.A. Crouch), even a 23% savings in identifying duplicates will be significant for assembling phenomic datasets going forward. Part of the reason for low precision is likely that characters are written in different languages or use very few terms (Table 1).

Richer, more verbose character descriptions should increase the chances of finding potential links between characters. As a final note, core principles of phylogenetic dataset construction remain as a key to phenomic analyses: character independence and mutually exclusive character states (Patterson, 1982). Addressing these issues in future dataset construction will limit the necessity of duplicate finding functionality in `phenotools` to evaluate the correlations among sets of characters.

Tools for visualizing phenomic data lag far behind those for genomic data. Going forward, we plan to enable compatibility with other data formats (e.g. NEXML) and improve the input of data matrices from PDF files (at present, this functionality is experimental and does not consistently work with diverse, non-tabular dataset formats). We feel there is a need for a robust phenome browser, similar to that available for genomic data (Deans et al., 2015). Desiderata for a prototype phenome browser include: the ability to compare phenomic differences to a reference species, options for highlighting differences in phenotypic traits among selected taxa, 'point-and-click' interaction tools for isolating linked sets of characters (e.g. through a web-based application), and 'clickable' characters that expand to a grid of photos of the labelled structures in different species. We hope that researchers working in image recognition will focus on challenges with automating the identification of bones or other phenotypic traits from images, which would constitute a tremendous advance in phenomics and bring us closer to the realization of a phenome for all life.

ACKNOWLEDGEMENTS

We thank T. Guillerme, D. Bapst, J. Proffitt, C. Torres, W. Gelnaw and L. English for insightful comments on earlier versions of this manuscript. Funding for this work was provided by the National Science Foundation (NSF EAR 1355292/DEB-1355343 to J.A.C and S.V.E.) and the Gordon and Betty Moore Foundation (4489) to J.A.C.

AUTHORS' CONTRIBUTIONS

J.A.C. and C.M.E. conceived the methodology. C.M.E. wrote the software and package documentation. All authors wrote the manuscript.

DATA ACCESSIBILITY STATEMENT

Data deposited in the Dryad repository: <https://doi:10.5061/dryad.05qm36k> (Eliason, C.M., Edwards, S.V., & Clarke, J.A, 2019b). Code is available on github (<https://github.com/celiason/phenotools>) and deposited at Zenodo (<https://doi.org/10.5281/zenodo.2650322>) (Eliason, C.M., Edwards, S.V., & Clarke, J.A, 2019a).

ORCID

Chad M. Eliason  <https://orcid.org/0000-0002-8426-0373>

Scott V. Edwards  <https://orcid.org/0000-0003-2535-6217>

Julia A. Clarke  <https://orcid.org/0000-0003-2218-2637>

REFERENCES

- Balhoff, J. P., Dahdul, W. M., Dececchi, T. A., Lapp, H., Mabee, P. M., & Vision, T. J. (2014). Annotation of phenotypic diversity: decoupling data curation and ontology curation using Phenex. *Journal of Biomedical Semantics*, 5(1), 1–5. <https://doi.org/10.1186/2041-1480-5-45>
- Baumel, J. J., & Witmer, L. M. (1993). Osteologia. In J. J. Baumel, A. S. King, J. E. Breazile, H. E. Evans, & J. C. Vanden Berge (Eds.), *Handbook of avian anatomy: Nomina anatomica avium* (Vol. 23). Cambridge: Publications of the Nuttall Ornithological Club.
- Bourdon, E., de Ricqlès, A., & Cubo, J. (2009). A new transantarctic relationship: Morphological evidence for a Rheidae-Dromaiidae-Casuariidae clade (Aves, Palaeognathae, Ratitae). *Zoological Journal of the Linnean Society*, 156(3), 641–663. <https://doi.org/10.1111/j.1096-3642.2008.00509.x>
- Brusatte, S. L., Lloyd, G. T., Wang, S. C., & Norell, M. A. (2014). Gradual assembly of avian body plan culminated in rapid rates of evolution across the dinosaur-bird transition. *Current Biology*, 24(20), 2386–2392. <https://doi.org/10.1016/j.cub.2014.08.034>
- Burleigh, G., Alphonse, K., Alverson, A. J., Bik, H. M., Blank, C., Cirranello, A. L., ... Yu, M. (2013). Next-generation phenomics for the Tree of Life. *PLoS Currents*, 24, <https://doi.org/10.1371/currents.tol.085c713acafc8711b2ff7010a4b03733>.
- Clarke, J. A., & Middleton, K. M. (2008). Mosaicism, modules, and the evolution of birds: Results from a Bayesian approach to the study of morphological evolution using discrete character data. *Systematic Biology*, 57(2), 185–201. <https://doi.org/10.1080/10635150802022231>
- Clarke, J. A., Zhou, Z., & Zhang, F. (2006). Insight into the evolution of avian flight from a new clade of Early Cretaceous ornithurines from China and the morphology of *Yixianornis grabaui*. *Journal of Anatomy*, 208, 287–308. <https://doi.org/10.1111/j.1469-7580.2006.00534.x>
- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695(5), 1–9.
- Cui, H. (2012). CharaParser for fine-grained semantic annotation of organism morphological descriptions. *Journal of the American Society for Information Science and Technology*, 63(4), 738–754. <https://doi.org/10.1002/asi.22618>
- Dahdul, W. M., Balhoff, J. P., Blackburn, D. C., Diehl, A. D., Haendel, M. A., Hall, B. K., ... Mabee, P. M. (2012). A unified anatomy ontology of the vertebrate skeletal system. *PLoS ONE*, 7(12), e51070. <https://doi.org/10.1371/journal.pone.0051070>
- Deans, A. R., Lewis, S. E., Huala, E., Anzaldo, S. S., Ashburner, M., Balhoff, J. P., ... Mabee, P. (2015). Finding our way through phenotypes. *PLoS Biology*, 13(1), e1002033. <https://doi.org/10.1371/journal.pbio.1002033>
- Eliason, C. M., Edwards, S. V., & Clarke, J. A. (2019a). Phenotools v. 1.0.0. <https://doi.org/10.5281/zenodo.2650322>
- Eliason, C. M., Edwards, S. V., & Clarke, J. A. (2019b). Data from: Phenotools: An R package for visualizing and analyzing phenomic datasets. *Dryad Digital Repository*, <https://doi.org/10.5061/dryad.05qm36k>
- Endara, L., Cui, H., & Burleigh, J. G. (2018). Extraction of phenotypic traits from taxonomic descriptions for the tree of life using natural language processing. *Applications in Plant Sciences*, 6(3), e1035. <https://doi.org/10.1002/aps.3.1035>
- Gavryushkina, A., Heath, T. A., Ksepka, D. T., Stadler, T., Welch, D., & Drummond, A. J. (2015). Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Systematic Biology*, 66(1), 57–73.
- Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B., Houde, P., Li, C., ... Zhang, G. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*, 346(6215), 1320–1331.
- Kemp, C. (2015). The endangered dead. *Nature*, 518, 292–294.
- Lee, K., Feinstein, J., & Cracraft, J. (1997). The phylogeny of ratite birds: resolving conflicts between molecular and morphological data sets.

- In David Mindell(Ed.), *Avian molecular evolution and systematics* (Vols 1–7, pp. 1–20).
- Lee, M., & Palci, A. (2015). Morphological phylogenetics in the genomic age. *Current Biology*, 25(19), R922–R929. <https://doi.org/10.1016/j.cub.2015.07.009>
- Liu, J., Endara, L., & Burleigh, J. G. (2015). MatrixConverter: Facilitating construction of phenomic character matrices. *Applications in Plant Sciences*, 3(2), 1400088. <https://doi.org/10.3732/apps.1400088>
- Maddison, W. P., & Maddison, D. R. (2018). Mesquite: a modular system for evolutionary analysis. Version 3.51. Retrieved from <http://www.mesquiteproject.org>
- McShea, D. W., & Venit, E. P. (2001). What is a part? In Günter P. Wagner (Ed.), *The character concept in evolutionary biology* (Vols 1–11, pp. 259–283). London: Academic Press.
- O’Leary, M. A., & Kaufman, S. (2011). MorphoBank: Phylophenomics in the “cloud”. *Cladistics*, 27(5), 529–537. <https://doi.org/10.1111/j.1096-0031.2011.00355.x>
- Patterson, C. (1982). Morphological characters and homology. In K. Josey, & A. Friday (Eds.), *Problems of phylogenetic reconstruction* (pp.21–74). London, UK:Academic Press.
- Ramírez, M. J., Coddington, J. A., Maddison, W. P., Midford, P. E., Prendini, L., Miller, J., ... Wheeler, W. C. (2007). Linking of Digital Images to Phylogenetic Data Matrices Using a Morphological Ontology. *Systematic Biology*, 56(2), 283–294. <https://doi.org/10.1080/10635150701313848>
- Schinke, R., Greengrass, M., Robertson, A. M., & Willett, P. (1996). A stemming algorithm for latin text databases. *Journal of Documentation*, 52(2), 172–187. <https://doi.org/10.1108/eb026966>
- van der Loo, M. (2014). The stringdist package for approximate string matching. *The R Journal*, 6(1), 111–122. <https://doi.org/10.32614/RJ-2014-011>
- Worthy, T. H., & Scofield, R. P. (2012). Twenty-first century advances in knowledge of the biology of moa (Aves: Dinornithiformes): A new morphological analysis and moa diagnoses revised. *New Zealand Journal of Zoology*, 39(2), 87–153. <https://doi.org/10.1080/03014223.2012.665060>
- Xu, X., You, H., Du, K., & Han, F. (2011). An Archaeopteryx-like theropod from China and the origin of Avialae. *Nature*, 475(7357), 465–470. <https://doi.org/10.1038/nature10288>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Eliason CM, Edwards SV, Clarke JA. `phenotools`: An R package for visualizing and analysing phenomic datasets. *Methods Ecol Evol*. 2019;10:1393–1400. <https://doi.org/10.1111/2041-210X.13217>