



All genera of the world: an overview and estimates based on the March 2020 release of the Interim Register of Marine and Nonmarine Genera (IRMNG)



TONY REES¹, LEEN VANDEPITTE^{2,3}, BART VANHOORNE^{2,4} & WIM DECOCK^{2,5}



¹ Private address, New South Wales, Australia.

 Tony.Rees@marinespecies.org;  <http://orcid.org/0000-0003-1887-5211>

² Flanders Marine Institute/Vlaams Instituut Voor De Zee (VLIZ), Wandelaarkaai 7, 8400 Ostend, Belgium.

³  leen.vandepitte@vliz.be;  <http://orcid.org/0000-0002-8160-7941>

⁴  bart.vanhoorne@vliz.be;  <https://orcid.org/0000-0002-6642-4725>

⁵  wim.decock@vliz.be;  <https://orcid.org/0000-0002-2168-9471>

Abstract

We give estimated counts of known accepted genera of the world (297,930±65,840, of which approximately 21% are fossil), of a total 492,620 genus names presently held for “all life”, based on the March 2020 release of the Interim Register of Marine and Nonmarine Genera (IRMNG). A further c. 9,400 accepted genus names are anticipated to have been published over the period 2014–2019 which are not yet included in IRMNG; together with a lower confidence estimate that perhaps an additional 3,000 historic names are also missing from the present version of IRMNG, we therefore estimate that approximately 310,000 accepted generic names have been published to the end of 2019, with the holdings of IRMNG being around 96% complete. A breakdown of the data is presented by phylum and, in some cases, lower taxonomic group such as class or order; the actual lists of names on which the totals are based are available for download via the IRMNG web site and are also included as supplementary data to this paper. These data provide the most complete and consistent coverage of all kingdoms of life presently available in such a form and, despite their “interim” nature (not completely vetted by taxonomic experts, not all genera yet placed to family), serve to illustrate the scope of a project for a more detailed survey of “all genera of the world” as well as providing a comparison with existing lists (for example, to indicate names that may be missed from either side), and preliminary content that can be of value for the compilation of new lists. We note areas (chiefly very recently published names) where present IRMNG data may be incomplete and briefly address other issues encountered in the assembly of such data, including those associated with the construction of a unified and/or consensus classification within which genera and their containing families can be placed.

Key words: taxonomy; systematics; taxonomic databases; genera; biodiversity informatics

Introduction

The concept of a series of papers addressing portions of the question of “all genera of the world” is a valuable one, which can benefit from as much preliminary scoping as may be currently available. To date, synoptic surveys of biodiversity have been attempted mainly at the level of family, for extant taxa (Parker 1982; Ruggiero 2014) or also with the added inclusion of (or primary focus upon) fossil taxa (Benton 1993; Zhang 2011, 2013). For species, the ongoing Catalogue of Life project (Roskov *et al.* 2019) continues to make progress, albeit without dedicated content relevant to genera; this is complemented in part by the Paleobiology Database (2020), which includes more detailed information on fossil genera but also suffers from a degree of incompleteness in its taxonomic coverage at the present time.

Extending a synoptic compilation from families to genera is an obvious next step and material for this exists in the form of the Interim Register of Marine and Nonmarine Genera (IRMNG), an online database initiated in 2006 (Rees 2008, 2020; Rees *et al.* 2017) with the aim of collating all published generic names into a single system, and in addition assigning “habitat” and “extant vs. fossil” indicators, the habitat flags initially set to either “marine” or “nonmarine” (or both), subsequently expanded into marine, brackish, freshwater and terrestrial (not yet complete for all names). The inclusion of the word “interim” in the compilation title is intended to convey that the data, while assembled mainly from “trusted sources”, are not completely vetted by taxonomic experts, as would be the case for a more authoritative compilation; nevertheless, bringing such data together in the form of an interim compilation still has benefits for users wishing to access a compendium that covers “all life” at the generic level. (Note, although we use “Register” in the IRMNG

title—in keeping with other similar projects such as the European Register of Marine Species (ERMS: Costello 2000) and the World Register of Marine Species (WoRMS Editorial Board 2020)—by contrast with “official” registers such as ZooBank, MycoBank, etc., IRMNG is not a formal register with any particular status or associated registration process, but merely a convenient assembly point for already available nomenclatural and taxonomic information.)

Compilation and maintenance of IRMNG has been greatly facilitated by the availability of pre-existing genus-level compendia, in particular the ten volumes of *Nomenclator Zoologicus* (Neave and successors 1939–2004) for animal names, the online *Index Nominum Genericorum* (Farr & Zijlstra 2020) for plants *sensu lato*, together with smaller compilations for prokaryotes (Parte 2020) and viruses (ICTV 2019). Bringing content from these sources (plus others) together provides a fairly complete coverage of “all life”, with additional, post-2004 animal names being available from the “Index to Organism Names” (Clarivate Analytics 2020) plus the primary literature. It should be noted that *Nomenclator Zoologicus* in particular, as its name indicates, is in essence a “nomenclator”, in other words a source of information regarding the authorship and place of publication of each included name, and not a taxonomic work concerned with the present taxonomic status and detailed placement names in a taxonomic hierarchy, so such taxonomic information must be sought elsewhere. Considering that zoological names make up the largest single component of all published names (this paper, Table 2 and Fig. 1), further researching these names represents a significant additional workload for the production of a compendium such as IRMNG, and one that is still incomplete for a subset of names (chiefly from this one source), as will be described further below.

With relevance to the series of papers to follow on the theme of “All genera of the world”, data from IRMNG can be useful in several respects. (1) Summary statistics from IRMNG can be used to indicate the scope and potential size of the problem to be addressed. (2) Where more completely taxonomically vetted compilations are already available for individual groups, comparison of those sources with equivalent content in IRMNG can lead to identification of data gaps (names missed) in either compilation, and/or assist with the ongoing improvement of IRMNG content to reflect the latest taxonomic opinions, for example in the recognition of valid (accepted) names versus synonyms, as well as higher taxonomic placement. (3) Where an “expert list” is in preparation, IRMNG content may be a useful source of material for consideration, in addition to other resources; and (4) where no expert list currently exists for a particular group (and may not be available for some time into the future), content from IRMNG can

provide an interim or proto-list at generic level to fill data gaps in other projects that desire to have a synoptic view of biodiversity, where genus-level detail is appropriate: examples of such projects presently include the Global Names Resolver (<https://resolver.globalnames.org/>), the Global Biodiversity Information Facility (GBIF) “backbone taxonomy” (<https://www.gbif.org/dataset/d7dddbf4-2cf0-4f39-9b2a-bb099caae36c>), the Open Tree of Life (<https://tree.opentreeoflife.org/taxonomy/browse>), and some areas of the Catalogue of Life, in particular those concerning selected Protista, Chromista, and algal groups currently treated as Plantae.

In the sections that follow, we describe aspects of IRMNG content relevant to the present topic, give estimates for total accepted genera in all taxonomic groups (all phyla plus certain classes and orders where these are of interest), and briefly address other relevant aspects of our experiences compiling and curating both IRMNG and the related World Register of Marine Species, which share a common data structure and are both now hosted at the Flanders Marine Institute (VLIZ) in Belgium (Vandepitte *et al.* 2015, 2018).

Selected characteristics of IRMNG content and the IRMNG data system (plus web portal)

IRMNG principle: one record per published name

The intention for IRMNG is to compile one record for every published name instance (new or replacement name, or in some cases, incorrect original or subsequent spellings, *nomina nuda*, etc.), with its authorship recorded as per established IRMNG conventions (Rees *et al.* 2017). Thus, IRMNG reconciles varying versions of the same “name instance” (for example, with different representations of what is the same set of authors) to a single record, in order that the number of published names can be accurately counted, and relevant characteristics such as taxonomic status and position, plus extant, habitat, and other attributes, can be associated with the name as applicable. Since IRMNG represents the set of all published names (subject to certain limitations as detailed below), and because this set includes both accepted names and synonyms plus other unaccepted names, the number of accepted taxa in any group will almost always be less than the number of published names.

Taxonomic status

Names stored in IRMNG have a “status” which can be set to values of accepted, unaccepted (in most cases with a link to the current accepted name), plus a small number of related taxonomic states, namely *nomen nudum*, *nomen dubium*, temporary name, *taxon inquirendum*, interim unpublished, or uncertain. For more details on the

usage of these states in both IRMNG and WoRMS we refer the reader to Horton *et al.* (2017). “Accepted” names (equivalent to valid names in zoology, current names in botany) and “unaccepted” names (plus related categories) are accorded these statuses based on external “trusted sources”; where recent such sources offer divergent opinions, an editorial decision is taken as to which source is preferred in individual cases, which can be accompanied by a relevant taxonomic note. “Uncertain” names are those which have been acquired from a nomenclator (such as *Nomenclator Zoologicus*) without indication of their present taxonomic status and have not yet been further investigated in that regard. As and when the “uncertain” names are further researched, they will end up being reassigned to either the “accepted” or one of the “unaccepted” or related categories.

“Candidatus” names in bacteriology (taxa in Bacteria and Archaea lacking one or more of the criteria for full acceptance) are presently included in the “accepted” category, although they have no official standing in prokaryotic nomenclature, since they are applied to taxa that are treated as current in the literature and for which associated published information is available, and are thus included in relevant taxon counts. (The number of such cases is small at present but could conceivably grow in future). As and when such names are replaced by validly published names according to the current prokaryotic Code of Nomenclature (Parker *et al.* 2019), the status of the name(s) in question can be reduced to a synonym of whatever new accepted name is eventually allocated.

Search options

Using the web entry point, IRMNG data can be searched and retrieved via both “basic” and “advanced” search interfaces. The basic interface permits search by all or any part of scientific name, authority, and IRMNG ID (vernacular names are not a part of present IRMNG content), and limited to either a search just on genera, or on names at any rank. Via the advanced interface, a web user can configure additional search options including to display fuzzy (approximate) as well as exact matches. The fuzzy search is useful to detect correctly spelled targets when a misspelled name is entered, as well as detect a range of similarly spelled names in the database in case these are of interest; this search employs the “Taxamatch” algorithm specifically developed for taxonomic names, for additional detail refer Rees (2014). The user can also limit the search results to, among other options:

- names of a particular rank and/or taxonomic status
- names in a particular taxonomic group (of major or intermediate rank, as held on the system)
- extant-only, fossil-only, marine-only names, etc., to the extent that these fields are populated in the

database

- names with a particular word or phrase in one of the “notes” fields
- names added, or edited, on the system within a specified date range and/or by a particular editor
- etc.

Common to all the databases hosted on the “Aphia” platform (including the WoRMS family of taxonomic databases), these options offer a very powerful way to customize searches to suit a user’s particular needs.

Options offered via additional links provided in the IRMNG web interface include a navigable Taxon tree, “Taxon match” and “Homonyms”. “Taxon match” (discussed further in Nozères *et al.* 2012 and Vandepitte *et al.* 2015) allows a web user to upload files of up to 1,000 taxonomic names at a time and generate a list of matching names complete with authorities, higher taxonomy, etc.; where homonyms exist, a drop-down selection box is supplied to allow the user to select the desired name instance prior to preparation of a data download where required (for example as a spreadsheet). If a user has the requirement to match more than 1,000 names, these can be submitted as a sequential set of batches, each under the pre-set limit; alternatively the entire database (principal fields only, but including the IRMNG ID as a primary key) is available as a data download, and can then be uploaded to a user’s own system (resources permitting) where unlimited and/or other custom queries can be run with a modest programming effort.

“Homonyms” takes the user to a series of pre-formatted links to lists of homonyms at both family and genus level. The latter are split into sub-lists alphabetically, since the number of homonyms at this rank means that without this, the lists would be very long. Such lists of homonyms are at present unique to IRMNG and are self-maintaining, in that as new names that are spelled the same as another already held (homonyms *sensu lato*) are added to the database, they will automatically generate a new entry on the list, and similarly an entry will disappear from the list if the number of instances of a name drops back to 1 (for example, if duplicate or erroneous entries are detected and removed).

Key recent sources used, and cut-off points for present IRMNG content

IRMNG is a continuous effort, with a degree of latency between names appearing in the published literature and their entry into the database, therefore it is relevant to note cut-off dates that apply to the totals presented below, with the expectation that the current totals as at the time of publication of this report are expected to be a little higher in most cases. Approximate cut-off dates relevant to the March 2020 release of IRMNG (used to generate the present counts) are as follows (Table 1):

TABLE 1. Latest sources and approximate cut-off dates for data in the present version of IRMNG by major taxonomic group.

Major group	Extant and/or fossil status	Latest sources used	Approximate cut-off date
Animals + zoological protists	extant and fossil	<ul style="list-style-type: none"> Index to Organism Names (ION) (Clarivate Analytics 2020, 2018 version) World Register of Marine Species (2016 version) 	end 2014 (ION data); mid 2016 (WoRMS data)
Land plants (bryophytes through angiosperms)	extant only	<ul style="list-style-type: none"> The Plant List version 1.1 (The Plant List 2013) International Plant Names Index (IPNI partnership 2020, 2018 version) 	2012 approx. (TPL data); end 2017 (IPNI data)
Fossil plants	fossil only	<ul style="list-style-type: none"> Taylor <i>et al.</i> (2009) Kansas University online Bibliography of Paleobotany (2009) <i>Index Nominum Genericorum</i> (2012 version) Novikoff & Barabasz-Krasny (2015) 	2009 (+) (note, later sources are not exhaustive)
Algae (includes selected Chromista, Protozoa, and “plant algae”)	mainly extant (a few taxa fossil)	<ul style="list-style-type: none"> <i>Index Nominum Genericorum</i> (2012 version) Relevant primary literature to 2016 WoRMS (October 2016 version) — includes some historic algal genus names from AlgaeBase not previously held 	mid 2016
Fungi	extant plus some fossil	<ul style="list-style-type: none"> Index Fungorum (Index Fungorum Partnership 2020) MycoBank (International Mycological Association 2020, August 2019 version) 	end 2013
Prokaryota (Archaea and Bacteria, including Cyanobacteria)	extant plus some fossil	<ul style="list-style-type: none"> List of Names with Standing in Prokaryotic Nomenclature (July 2018 version) CyanoDB (Hauer & Komárek 2020, January 2019 version) Raaben <i>et al.</i> (2001) (fossil stromatolites) 	mid 2018 (LPSN and CyanoDB)
Viruses	extant only	<ul style="list-style-type: none"> the ICTV database (March 2018 release, ratified 2019) 	end 2017
Pseudofossils	fossil only	<ul style="list-style-type: none"> Häntzschel (1975), plus some more recent literature 	1975 (+)

Known residual gaps are mainly in the area of fossil plant genera known to exist via literature summaries (e.g. Watt 1982, Schultze-Motel 2003) but in some cases absent from *Index Nominum Genericorum*, together with other recently published fossil plant names (including dispersed spores and pollen) not yet sought in the primary literature.

As at March 2020, IRMNG contains 492,620 genus names of which 297,930 are estimated to be accepted (actual range 232,090–363,770 depending upon whether or not “uncertain” names are included). This leaves a residue of 194,690 names (a combined total of 128,850 known unaccepted names plus a notional 50% of the 131,680 “uncertain” names) which are likely to be synonyms, misspellings, or otherwise unavailable names; where known, such names are then pointed in IRMNG to their equivalent accepted name, although for names in the

“uncertain” category, this process will be dependent upon additional work.

Taxonomic arrangement

Higher taxonomy in IRMNG follows a single (unified) scheme, constructed so far as is possible to be congruent with a recent “consensus classification”, that of Ruggiero *et al.* (2015) being used as the favoured example at this time for most taxa (exceptions are given below). The sometimes different approach by Adl *et al.* (2019) is not used, partly because these authors present their hierarchy without strict Linnean ranks, and partly because moving to that system would remove concordance between IRMNG and the treatment currently employed in the Catalogue of Life, which also follows Ruggiero *et al.* Where the Ruggiero *et al.* (2015) treatment is already superseded, or in one case (the treatment of Aves as below) is preferred

not to be followed, differences can exist which are presently as follows:

- Aves is retained at class level within Chordata, in accordance the most recent (2019, 2020) preferences of the Catalogue of Life and WoRMS; this group is considered a subclass of Reptilia in Ruggiero *et al.* (2015)
- the treatment of fishes in IRMNG is awaiting revision. The present IRMNG classes Actinopterygii (ray-finned fishes), Chondrichthyes (cartilaginous fishes) and Sarcopterygii (lobe-finned fishes) are treated as superclasses by Ruggiero *et al.* (2015), which, however, does not include the numerous extinct groups. The main alternative published treatment, that of Nelson *et al.* (2016), does include extinct forms but differs in some significant respects from the Ruggiero *et al.* (2015) treatment, for example recognising Chondrichthyes as a class and Actinopterygii and Sarcopterygii as subclasses of Osteichthyes, a class not recognised by Ruggiero *et al.*; editorial decisions as to how best combine elements of both systems for IRMNG purposes are yet to be made. IRMNG also includes some (“uncertain”) genus names presently allocated just to “Pisces”; the latter is no longer a formal taxonomic group but is the designation for these names in *Nomenclator Zoologicus*, which they retain in IRMNG until further reviewed
- the protozoan phylum Hemimastigophora is reinstated based on Lax *et al.* (2018)
- a new algal phylum Rhodelphidia (sister to Rhodophyta) is introduced in Plantae as per Gawryluk *et al.* (2019)
- the phylum Kinorhyncha is reorganised following Sørensen *et al.* (2015)
- various portions of the Chromistan and Protozoan classification have been upgraded following the most recent publications by Cavalier-Smith and co-workers (Cavalier-Smith 2016; Cavalier-Smith *et al.* 2015, 2016, 2018)
- a new protozoan phylum Aphelida has been added as per Karpov *et al.* (2014) (previously a class); Adl *et al.* (2019) also recognise this group (under the name Aphelidea), but include it within Fungi
- a new phylum (Entomophthoromycota) and various new classes and orders of Fungi have been introduced post the treatment of Ruggiero *et al.* (2015), as per records in MycoBank, *Index Fungorum*, and the primary literature.

For land plants, a treatment has been developed for IRMNG which merges the system of Novikoff & Barabasz-Krasny (2015) for fossils with that for extant plants given by Ruggiero *et al.* (2015). To achieve this, several of Novikoff & Barabasz-Krasny’s phyla (“divisions”) have been reduced in rank, for example their divisions Rhyniophyta, Zosterophyllophyta, Progymnospermophyta and Pteridospermophyta are treated as classes (Rhyniopsida, etc.) within phylum Tracheophyta in IRMNG, while their division Langiophytrophyta is treated as the earlier phylum name Horneophyta. Their divisions Cycadophyta, Ginkgophyta and Gnetophyta are also treated as classes (not phyla) within Tracheophyta, within superclass Gymnospermae.

In the case of certain groups of sometimes doubtful taxonomic position, editorial decisions have been taken as to where they should presently be placed in the IRMNG hierarchy; such decisions can always be revisited if and when relevant new information is available. Specifically:

- Microsporidia are treated as a phylum within Sarcomastigota (Protozoa) as per the treatment of Ruggiero *et al.* (2015); these, along with similar small groups such as rozellids, are alternatively included as taxa of lower rank within basal Fungi in some treatments, e.g. those of Tedersoo *et al.* (2018) and Adl *et al.* (2019)
- Acritarcha and Calcitarcha (organic-walled and calcareous cyst-like microfossils of unknown affinities) are placed in Protozoa, although some or many are probably algae—whether Chromista or Plantae is not known
- Chitinozoa (flask-shaped, organic walled marine microfossils) are placed in Animalia, since they have been hypothesised to be either eggs or juvenile stages of an unknown type of animal, although they have also been suggested to have protist, plant or fungal affinity; for additional information refer Jain (2020)
- *Sporae dispersae* (fossil dispersed spores and pollen) are placed as their own group (“phylum”) in Plantae; most are probably from Tracheophyta (Angiospermae and Gymnospermae), though some may represent spores of Ferns, Bryophyta, Fungi or other early land plants
- Receptaculitids are treated as fossil algae (as per Rietschel & Nitecki 1984), not sponges or Problematica, and are placed as an order in Chlorophyceae since other workers consider them allied to, or within, the Dasycladales (Rietschel & Nitecki treat them in their own algal class, Receptaculitaphyceae)
- Archaeocyatha, listed as a separate phylum in IRMNG prior to this release, have been re-ranked

as an extinct class of phylum Porifera (sponges) as per most current treatments, refer e.g. Kerner *et al.* (2011). This adds around 550 genus names in total (308 accepted) to “Porifera” as previously recognised (representing a slight increase in the proportion of fossil taxa), and reduces the number of recognised phyla in IRMNG accordingly

- The phylum Lobopoda was introduced by Cavalier-Smith (1998) as a new phylum to include (extant) onychophorans and tardigrades, the latter being then reduced to subphyla. In IRMNG, onychophorans and tardigrades are retained at phylum level as per the treatment of Ruggiero *et al.* (2015); the phylum name “Lobopoda” (in addition to Onychophora and Tardigrada) is kept for fossil, stem-group “lobopods” which are not presently assignable to either of the latter two phyla
- The fossil phylum Vendobionta (considered as partially equivalent to or overlapping phylum Proarticulata Fedonkin) is reorganised into 3 classes in accordance with the treatment of Ivantsov *et al.* (2019)
- Separate categories (currently treated at phylum rank) are retained in IRMNG for both “Trace fossils (unallocated)” and “Problematica”. Where a trace fossil is unambiguously assigned to an organism of a particular taxon that is considered to have created it (for example Annelida, Mollusca, Reptilia, etc.) it is removed to that group, with the exception of fossil coprolites of invertebrates, a single ichnofamily of which can contain forms attributed to a range of producers (Knaust 2020). “Problematica” contains a variety of fossil taxa at different ranks ranging from class to genus which appear to be animal in nature but whose exact relationships are unclear at this time, and have previously been referred to using this terminology
- Fossil stromatolites (biogenic mounds attributed to the action of Cyanobacteria) are assigned to Cyanobacteria, although they are treated as trace fossils (using zoological nomenclature) in some other systems. Since the main treatment followed in IRMNG (that of Raaben *et al.* 2001) uses zoological nomenclature for families names in this group, these names (with their zoological endings) are retained in IRMNG despite the included taxa being listed with the Cyanobacteria, which otherwise follow the botanical nomenclatural Code
- Two miscellaneous “algal” categories (treated as phyla), “Algae *incertae sedis*” and “Algae (awaiting allocation)”, are presently placed

in kingdom Plantae where the green and red algae reside, although some may belong in Chromista (brown algae, diatoms, etc.), Protozoa (euglenoids), or Cyanobacteria (blue-green algae)

- Subviral agents (including prions, satellites and viroids) are presently included as separate “classes” within their own “phylum” in kingdom Viruses, although not all of these groups are included in the latest ICTV classifications
- Names of taxa currently believed to be pseudofossils (non biotic) and/or of questionable biological affinity have been removed from Animalia and Plantae, and are placed in their own group “Questionable/non-biota (fossil)”, currently treated at an equivalent level (kingdom). This contains no accepted taxa, but the names remain available under the rules of nomenclature for purposes of homonymy, etc., and can be reinstated and moved elsewhere if the taxa concerned are reassessed as indeed being legitimate biological entities.

Extant vs. fossil genera in IRMNG

IRMNG records can be flagged as Recent (=extant) only, Recent+fossil, fossil only, or unknown (=not yet researched). Since most fossil records are believed to be flagged as such, it is presumed that most of the “unknown” records will equate to “Recent” in due course. For data summary purposes, only records flagged “fossil only” are included in the fossil proportion quoted below, all others being treated as extant.

Procedure for estimating totals in this report

To produce estimates for this paper, for each group we cite a “low value” comprising the present accepted total only, a “high value” which comprises both the “accepted” and “uncertain” names combined, and the mean of the two together with an associated range, the latter being 50% of the total “uncertain” names; we believe that these values represent the best estimates that can be produced based on present data. The actual lists of names on which the totals are based are supplied as supplementary information to this paper, and can also be downloaded on demand from the IRMNG website (for details see “Data availability”).

Results

Genus totals by taxonomic group

Estimated numbers / percentages of accepted genera
(totals as per Table 2) - by kingdom

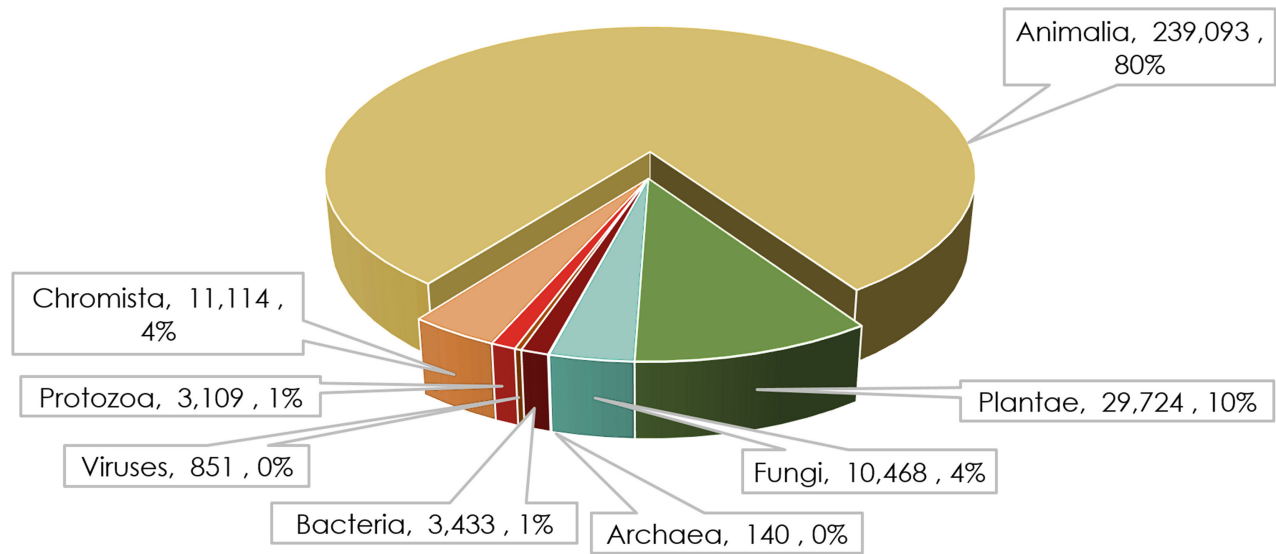


FIGURE 1. Kingdom-level breakdown of IRMNG content (March 2020 version) by estimated numbers of accepted genera. Note, “Algae” of older treatments no longer appear as a kingdom, being now distributed among Plantae, Chromista and Protozoa, with former “blue-green algae” (Cyanobacteria) now within Bacteria. Kingdom Chromista of this scheme (based on Ruggiero *et al.* 2015) corresponds approximately to the “SAR” (or Sar) supergroup (Stramenopiles, Alveolates and Rhizaria) of other schemes e.g. that of Adl *et al.* (2019), with the addition of Cryptista, Haptophyta, Picozoa and Telonemia.

Detailed breakdown by taxonomic group

TABLE 2. Genus holdings in IRMNG, March 2020 version.

Taxonomic group	all genera	accepted genera: low limit (= actual)	accepted genera: high limit (= actual+ uncertain)	accepted genera: mean value	range (±)	% genera fossil (accepted only)
All groups	492,620	232,090	363,770	297,930	65,840	21.4%
..Kingdom Chromista	15,360	9,846	12,382	11,114	1,268	39.1%
....Phylum Acavomonidia	1	1	1	1	0	0%
....Phylum Bigyra	109	62	84	73	11	0%
....Phylum Cercozoa	362	275	302	289	14	6.2%
....Phylum Ciliophora	2,169	1,596	1,888	1,742	146	2.2%
....Phylum Cryptista	65	47	50	49	2	0%
....Phylum Haptophyta	690	430	544	487	57	71.9%
....Phylum Heliozoa	30	22	23	23	1	0%
....Phylum Miozoa	2,032	1,412	1,540	1,476	64	45.4%
....Phylum Ochrophyta	2,554	1,894	1,988	1,941	47	18.5%
....Phylum Picozoa	1	1	1	1	0	0%
....Phylum Pseudofungi	199	140	140	140	0	7.1%
....Phylum Retaria	7,145	3,963	5,818	4,891	928	62.8%

...Continued on the next page

TABLE 2. (Continued)

Taxonomic group	all genera	accepted genera: low limit (= actual)	accepted genera: high limit (= actual+ uncertain)	accepted genera: mean value	range (±)	% genera fossil (accepted only)
....Phylum Rhizaria <i>incertae sedis</i>	1	1	1	1	0	0%
....Phylum Telonemia	1	1	1	1	0	0%
..Kingdom Protozoa	5,045	1,903	4,314	3,109	1,206	45.1%
....Phylum Acritarcha	1,005	737	945	841	104	100%
....Phylum Amoebozoa	524	321	363	342	21	0.6%
....Phylum Aphelida	3	3	3	3	0	0%
....Phylum Calcitarcha	57	53	53	53	0	100%
....Phylum Choanozoa	153	108	114	111	3	0%
....Phylum Euglenozoa	302	162	218	190	28	2.5%
....Phylum Hemimastigophora	4	3	3	3	0	0%
....Phylum Loucozoa	8	7	7	7	0	0%
....Phylum Metamonada	236	156	180	168	12	8.3%
....Phylum Microsporidia	242	197	200	199	2	0%
....Phylum Percolozoa	21	17	17	17	0	0%
....Phylum Protozoa (awaiting allocation)	1,727	35	1,484	760	725	5.7%
....Phylum Protozoa <i>incertae sedis</i>	78	63	66	65	2	69.8%
....Phylum Sarcomastigota (awaiting allocation)	664	23	643	333	310	17.4%
....Phylum Sulcozoa	21	18	18	18	0	0%
..Kingdom Animalia	393,235	183,743	294,442	239,093	55,350	23.0%
....Phylum Acanthocephala	297	160	178	169	9	0%
....Phylum Agmata	4	3	3	3	0	100%
....Phylum Animalia (awaiting allocation)	283	19	226	123	104	100%
....Phylum Animalia <i>incertae sedis</i>	3	2	2	2	0	100%
....Phylum Annelida	5,371	2,380	3,423	2,902	522	12.7%
....Phylum Arthropoda	234,458	118,537	184,856	151,697	33,160	9.4%
.....Subphylum Chelicerata	21,252	13,591	15,309	14,450	859	7.0%
.....Subphylum Crustacea	21,140	13,507	17,121	15,314	1,807	25.4%
.....Subphylum Hexapoda	181,943	87,234	143,191	115,213	27,979	4.3%
.....Class Collembola	1,037	285	934	610	325	3.9%
.....Class Diplura	197	139	140	140	1	3.6%
.....Class Insecta	180,621	86,733	142,040	114,387	27,654	4.3%
.....Class Protura	88	77	77	0	0	1.3%
.....Subphylum Myriapoda	4,040	1,165	3,637	2,401	1,236	3.4%
.....Class Chilopoda	723	166	498	332	166	6.6%
.....Class Diplopoda	3,220	945	3,055	2,000	1,055	2.4%
.....Class Pauropoda	71	39	63	51	12	0%
.....Class Symphyla	19	9	14	12	3	0%

...Continued on the next page

TABLE 2. (Continued)

Taxonomic group	all genera	accepted genera: low limit (= actual)	accepted genera: high limit (= actual+ uncertain)	accepted genera: mean value	range (±)	% genera fossil (accepted only)
.....other Myriapoda	7	6	7	7	1	100%
.....other Arthropoda	6,083	3,040	5,598	4,319	1,279	100%
....Phylum Brachiopoda	6,531	4,628	5,697	5,163	535	96.8%
....Phylum Bryozoa	3,540	2,197	3,024	2,611	414	58.5%
....Phylum Cephalorhyncha <i>incertae sedis</i>	1	1	1	1	0	100%
....Phylum Chaetognatha	74	40	54	47	7	22.5%
....Phylum Chitinozoa	117	56	116	86	30	100%
....Phylum Chordata	60,084	20,753	37,933	29,343	8,590	43.1%
.....Class Amphibia	2,334	739	1,844	1,292	553	21.1%
.....Class Aves	13,481	2,589	7,951	5,270	2,681	10.9%
.....Class Mammalia	13,609	5,332	8,952	7,142	1,810	75.9%
.....Class Reptilia	9,339	3,973	7,459	5,716	1,743	48.9%
.....other Chordata (“Pisces”, etc.)	21,321	8,120	11,727	9,924	1,804	31.0%
....Phylum Cnidaria	9,308	3,780	6,886	5,333	1,553	51.5%
....Phylum Ctenophora	123	72	83	78	6	16.7%
....Phylum Cycliophora	1	1	1	1	0	0%
....Phylum Cycloneuralia <i>incertae sedis</i>	2	2	2	2	0	100%
....Phylum Echinodermata	7,897	4,031	5,766	4,899	868	66.5%
....Phylum Entoprocta	28	14	15	15	1	0%
....Phylum Gastrotricha	120	66	75	71	5	0%
....Phylum Gnathostomulida	29	27	28	28	1	0%
....Phylum Hemichordata	679	336	582	459	123	90.8%
....Phylum Hyolitha	203	125	192	159	34	100%
....Phylum Kinorhyncha	39	26	26	26	0	0%
....Phylum Lobopoda	21	18	20	19	1	100%
....Phylum Loricifera	11	11	11	11	0	0%
....Phylum Micrognathozoa	1	1	1	1	0	0%
....Phylum Mollusca	42,644	16,069	30,289	23,179	7,110	55.4%
....Phylum Nematoda	5,451	3,163	3,323	3,243	80	3.2%
....Phylum Nematomorpha	37	22	23	23	1	9.1%
....Phylum Nemertea	508	348	393	371	23	1.1%
....Phylum Onychophora	67	48	55	52	4	6.3%
....Phylum Orthonectida	7	6	6	6	0	0%
....Phylum Phoronida	13	5	5	5	0	60%
....Phylum Placozoa	4	3	3	3	0	0%
....Phylum Platyhelminthes	7,592	3,823	5,411	4,617	794	1.8%
....Phylum Porifera	5,435	2,197	3,856	3,027	830	64.6%
....Phylum Priapula	87	69	71	70	1	88.4%
....Phylum Problematica	1,057	169	1,033	601	432	100%

...Continued on the next page

TABLE 2. (Continued)

Taxonomic group	all genera	accepted genera: low limit (= actual)	accepted genera: high limit (= actual+ uncertain)	accepted genera: mean value	range (±)	% genera fossil (accepted only)
....Phylum Rhombozoa	11	9	9	9	0	0%
....Phylum Rotifera	412	126	181	154	28	0%
....Phylum Scalidophora <i>incertae sedis</i>	1	1	1	1	0	100%
....Phylum Sipuncula	116	22	33	28	6	18.2%
....Phylum Tardigrada	156	132	144	138	6	3.0%
....Phylum Trace fossils (awaiting allocation)	367	184	319	252	68	100%
....Phylum Vendobionta	81	54	73	64	10	100%
....Phylum Vetulicolia	14	11	13	12	1	100%
....Phylum Vinctiplicata <i>incertae sedis</i>	3	3	3	3	0	100%
....Phylum Xenacoelomorpha	150	118	120	119	1	1.7%
..Kingdom Archaea	157	140	140	140	0	0%
....Phylum Crenarchaeota	35	31	31	31	0	0%
....Phylum Euryarchaeota	116	106	106	106	0	0%
....Phylum Korarchaeota	2	1	1	1	0	0%
....Phylum Nanoarchaeota	1	0	0	0	0	n/a
....Phylum Thaumarchaeota	3	2	2	2	0	0%
..Kingdom Bacteria	4,198	3,318	3,548	3,433	115	6.8%
....Phylum Acidobacteria	25	25	25	25	0	0%
....Phylum Actinobacteria	476	413	413	413	0	0.2%
....Phylum Aquificae	15	15	15	15	0	0%
....Phylum Armatimonadetes	3	3	3	3	0	0%
....Phylum Bacteria (awaiting allocation)	77	7	13	10	3	42.9%
....Phylum Bacteria <i>incertae sedis</i>	29	11	11	11	0	0%
....Phylum Bacteroidetes	388	365	365	365	0	0%
....Phylum Balneolaeota	4	4	4	4	0	0%
....Phylum Caldiserica	1	1	1	1	0	0%
....Phylum Calditrichaeota	2	2	2	2	0	0%
....Phylum Chlamydiae	9	6	6	6	0	0%
....Phylum Chlorobi	16	10	10	10	0	0%
....Phylum Chloroflexi	28	26	26	26	0	0%
....Phylum Chrysiogenetes	3	3	3	3	0	0%
....Phylum Cyanobacteria	1,068	593	815	704	111	36.9%
....Phylum Deferribacteres	7	7	7	7	0	0%
....Phylum Deinococcus- Thermus	10	10	10	10	0	0%
....Phylum Dictyoglomi	1	1	1	1	0	0%
....Phylum Elusimicrobia	1	1	1	1	0	0%
....Phylum Fibrobacteres	3	3	3	3	0	0%

...Continued on the next page

TABLE 2. (Continued)

Taxonomic group	all genera	accepted genera: low limit (= actual)	accepted genera: high limit (= actual+ uncertain)	accepted genera: mean value	range (±)	% genera fossil (accepted only)
....Phylum Firmicutes	552	485	485	485	0	0%
....Phylum Fusobacteria	12	11	11	11	0	0%
....Phylum Gemmatimonadetes	2	2	2	2	0	0%
....Phylum Kiritimatiellaeota	1	1	1	1	0	0%
....Phylum Lentisphaerae	3	3	3	3	0	0%
....Phylum Nitrospira	4	3	3	3	0	0%
....Phylum Planctomycetes	29	24	24	24	0	0%
....Phylum Proteobacteria	1,314	1,191	1,191	1,191	0	0.1%
....Phylum Rhodothermaeota	7	7	7	7	0	0%
....Phylum Spirochaetae	28	21	21	21	0	0%
....Phylum Synergistetes	15	15	15	15	0	0%
....Phylum Tenericutes	24	11	11	11	0	0%
....Phylum Thermodesulfobacteria	5	4	4	4	0	0%
....Phylum Thermotogae	13	12	12	12	0	0%
....Phylum Verrucomicrobia	23	22	22	22	0	0%
..Kingdom Fungi	16,932	10,286	10,649	10,468	182	4.0%
....Phylum Ascomycota	12,253	7,496	7,612	7,554	58	1.5%
....Phylum Basidiomycota	3,387	1,902	1,941	1,922	20	1.4%
....Phylum Blastocladiomycota	24	15	15	15	0	6.7%
....Phylum Chytridiomycota	203	161	163	162	1	6.8%
....Phylum Entomophthoromycota	36	23	24	24	1	0%
....Phylum Fungi (awaiting allocation)	656	447	623	535	88	53.5%
....Phylum Fungi <i>incertae sedis</i>	50	19	44	32	13	52.6%
....Phylum Glomeromycota	55	40	40	40	0	12.5%
....Phylum Zygomycota	268	183	187	185	2	6.0%
..Kingdom Plantae	56,614	22,003	37,444	29,724	7,721	9.8%
....Phylum Algae (awaiting allocation)	376	63	309	186	123	79.4%
....Phylum Algae <i>incertae sedis</i>	88	43	53	48	5	97.7%
....Phylum Anthocerotophyta	22	13	16	15	2	15.4%
....Phylum Bryophyta	1,641	997	1,377	1,187	190	1.8%
....Phylum Charophyta	361	192	272	232	40	42.2%
....Phylum Chlorophyta	1,969	1,144	1,518	1,331	187	18.5%
....Phylum Glaucophyta	8	6	6	6	0	0%
....Phylum Horneophyta	11	10	10	10	0	100%
....Phylum Marchantiophyta	830	281	642	462	181	4.6%
....Phylum Nematophyta	11	9	9	9	0	100%

...Continued on the next page

TABLE 2. (Continued)

Taxonomic group	all genera	accepted genera: low limit (= actual)	accepted genera: high limit (= actual+ uncertain)	accepted genera: mean value	range (±)	% genera fossil (accepted only)
....Phylum Plantae (awaiting allocation)	911	60	790	425	365	100%
....Phylum Rhodelphidia	1	1	1	1	0	0%
....Phylum Rhodophyta	1,711	1,015	1,015	1,015	0	6.9%
....Phylum Sporae dispersae	2,885	312	2,679	1,496	1,184	100%
....Phylum Tracheophyta	45,789	17,857	28,614	23,236	5,379	7.1%
.....Class Polypodiopsida	1,895	633	1,058	846	213	47.2%
.....Superclass Angiospermae	41,388	16,351	25,338	20,845	4,494	1.2%
.....Superclass Gymnospermae	1,834	594	1,562	1,078	484	85.2%
.....other Tracheophyta	672	279	656	468	189	96.8%
..Kingdom Questionable/non-biota (fossil)	90	0	0	0	0	n/a
..Kingdom Viruses	989	851	851	851	0	0%
....Phylum Subviral agents	21	21	21	21	0	0%
....Phylum Viruses	968	830	830	830	0	0%

Genus totals by year of publication (all groups except viruses)—period 1980–2019

All genera in IRMNG, with the exception of viruses plus a few records imported from early sources utilised, are accompanied by their year of publication (in botany as well as in zoology), which permits analysis of IRMNG holdings according to the year in which each name was published. This can be of value in studying trends in publication rates over time and, when used in a predictive manner, can provide the basis for estimating numbers of recently published names not yet represented in IRMNG. Relevant totals for the current version of IRMNG are presented in Table 3.

From Table 3 it can be seen that both the annual number of newly published names as recorded in IRMNG, and the proportion of these considered accepted, has remained relatively constant over the period 1990–2013, at around 2,500 published (2,200 accepted) names per year. Using the data from 2010–2013 (inclusive) as a guide (mean value 2,241 new, accepted genera per year), and extrapolating for the period 2014–2019 (inclusive) there would be an expected 13,446 new, accepted genus descriptions over this period, over which IRMNG presently holds 4,039, a projected deficit of around 9,400 names. It is therefore reasonable to presume that the estimated total reported here for IRMNG mean accepted genera for “all life” (297,930) will need to be adjusted upwards by around this value (9,400 plus an additional, very small component for

new virus genera added since 2017), plus an allowance for any historic names missed (perhaps 3,000: refer footnote to Table 4) in order to arrive at a true estimate of accepted, published generic names to end 2019.

Discussion

Comparisons with previous work

To the authors’ knowledge, this work represents the first attempt to provide estimates of global biodiversity by major taxon at the generic level, either for all published names, or for accepted-only names, i.e. nominal taxa, with additional reporting of the extent of the fossil-only component for each group (noting, for example, that some groups are 100% fossil). Chapman (2009) produced somewhat similar tables for estimated numbers of extant species, both described and undescribed, in major groups, however his data were based on a combination of both published databases and on expert estimates, of which the latter can be prone to conscious or unconscious biases (Costello 2015) and cannot be comprehensively checked since compilations of all the relevant source data do not exist.

From the standpoint of data verifiability, the long running Catalogue of Life (CoL) project, currently in its nineteenth annual release, offers a more reliable prospect at species level, although its treatment of genera is minimal other than supplying a simple name in each case, and its

TABLE 3. Counts of all, and accepted genus names in IRMNG by year of publication, 1980–2019 (excluding viruses).

Publication year	IRMNG content: generic names published/year—all names (including subgenera in zoology; viruses omitted) (^ = data notionally incomplete)	Same—accepted genus names only (low and high values as per Table 1) (^ = data notionally incomplete)
1980–1989 (yearly average)	2,943	low: 1,617; high: 2,619; mean: 2,118
1990–1999 (yearly average)	2,480	low: 1,285; high: 2,290; mean: 1,788
2000–2009 (yearly average)	2,589	low: 1,863; high: 2,375; mean: 2,119
2010	2,513	2,351
2011	2,270	2,107
2012	2,509	2,189
2013	2,642	2,318
2014	2,217^	1,987^
2015	675^	640^
2016	592^	569^
2017	479^	465^
2018	199^	188^
2019	195^	190^

coverage, while steadily improving, is still incomplete for some groups including Acari, Nematoda, Platyhelminthes, some Protozoa and many Chromista and “plant algae”; values for the latter are available separately via AlgaeBase (Guiry & Guiry 2020). The CoL does provide summary statistics of its present holdings of accepted, living species per taxonomic group (as compared in many cases against relevant “expert estimates”) in order to calculate percentage completeness, via the page at <http://www.catalogueoflife.org/annual-checklist/2019/info/totals>. However, equivalent statistics are not presently provided for genera, and genus names in the CoL are not associated with either authors or publication years¹.

The Index to Organism Names (ION) compilation publishes summary statistics—available at <http://www.organismnames.com/metrics.htm?page=graphs>—on newly published zoological genera and subgenera (combined) that have been encountered by their literature

searches conducted as inputs to “Zoological Record”, without discriminating extant from fossil names, genera from subgenera, and accepted names from synonyms. ION data will be more complete than IRMNG for the years 2015–2019 (no 2020 names are yet reported) and as an example, over the years 2010–2014 inclusive, the totals given are 1,897/1,908/1,993/1,979/1,891 while totals for 2015–2019 are 1,935/1,864/1,622/1,850/1,303 (data as at 6 March 2020)². Disregarding the 2019 value as not yet complete, this gives a mean value of “all names” (in zoology only) of 1,934 for 2010–2014 and 1,818 for 2015–2018 and indicate first that the 2010–2014 totals for IRMNG given in the first column of Table 3 are reasonable (in the order of 2,200–2,500 new names per year for “all life”), and second, that the presumption of rates of description continuing at a more or less constant level in more recent years (used for estimating the potential value for names missed) is generally realistic.

Unavailable and/or invalidly published names

In zoology, names are divided into available and unavailable, the latter set including *nomina nuda*, original and subsequent misspellings (but not emendations), some suppressed names, and other names that do not meet the

¹ Devine & Coddington (2019) present a totals for higher taxa from kingdom through genus in the 2019 release of the Catalogue of Life, including a value of 165,683 for genera, about half the estimated total given here for “all life”. The reasons for this discrepancy are yet to be fully explored but doubtless include the present incompleteness of the CoL (presently claimed to be around 80% complete for extant species only) and the lack of extensive CoL coverage of fossils. It is also possible that some generic names in IRMNG presently listed as “accepted” may turn out to incorrectly assigned (refer Table 4 for additional discussion), that a higher proportion than 50% of the present “uncertain” names may represent unaccepted rather than potentially accepted names, or that the COL is less complete across all groups than is currently stated.

² It is regrettable—although perhaps understandable in view of the fact that Clarivate Analytics is a commercial operation, and may wish to restrict certain services only to their paid subscribers—that the public user cannot then proceed to obtain the lists of “new names by year” directly via the ION website. Such data would make a useful comparison with IRMNG and might, for example, enable the detection and subsequent upload of a small number of additional names missed by the present IRMNG ingestion process.

general conditions for availability (ICZN 1999, 2012); in botany, equivalent categories are “validly published” names versus names not validly published (Turland *et al.* 2018). Unavailable and/or invalidly published names (including misspellings) do not exist for nomenclatural purposes under the relevant Codes and do not enter homonymy or synonymy, however, since a number of these are included in *Nomenclator Zoologicus* and other published sources, the decision has been made to retain them in IRMNG since they do occur with other associated data of interest such as specimen or locality information, and can also be useful for query expansion and taxonomic name resolution (for additional discussion refer Boyle *et al.* 2013).

Accuracy of estimates derived from IRMNG

The values included in this report are believed to be the best estimates available based on the present database content, but could be affected by some residual errors and missing data as reflected in the “interim” nature of the IRMNG compilation. Some issues of potential significance can be itemised as follows (Table 4):

General caveats on IRMNG data

As indicated above, IRMNG is presently offered as an “interim” product where (to a certain extent) depth (as in, for example, the degree of additional research for every entry) has been traded for breadth (an initial attempt to gather “all names”, with or without the full taxonomic information that might ideally be desired); the latter can, nevertheless, be added via subsequent passes over relevant content. To this degree, a certain level of incompleteness will be encountered by the user wishing to obtain (for example) complete lists by family, or a final total of accepted names for any group. In addition, the database may contain errors—either uncorrected errors imported with source data, or introduced via some subsequent process—which can, however, be corrected wherever discovered, either by the system compiler(s) or as notified to the custodians by external users. We therefore caution that IRMNG data may be less up-to-date than the current literature, to which reference should be made for the most “expert assessed” recent content.

Also as mentioned above, IRMNG-generated lists by family may be incomplete, although in the majority of cases relevant names will still be held at a higher taxonomic placement such as “Mammalia (awaiting allocation)”,

TABLE 4. Selected potential IRMNG issues and their possible effects.

Issue	Effect
A small percentage of historic names are not yet held (e.g. if missing in major nomenclators, and not encountered in other sources used to date)	Most are probably older synonyms etc. (these would go largely into “unaccepted” total), except in the case of some known missing fossil plant names and <i>Sporae dispersae</i> . For some additional considerations refer footnote ³
Around 9,400 recently published names are estimated to be missing (refer discussion of data in Table 3), plus a small number of unaccepted names to the end of 2019)	Cited totals will be lower than actual numbers of names published for the period 2014 (approx.) to 2019
Some accepted names may be erroneously flagged unaccepted (or vice versa)	“Accepted” totals would be under- or overestimates
Some extant names may be erroneously flagged fossil (or vice versa)	“Extant” and “fossil” name totals would be under- or overestimates
A small proportion of “uncertain” names may turn out to be non-names (e.g. database errors) or duplicates upon further investigation	“Uncertain” totals may be slight overestimates (anticipated effect on reported totals would be very minor)
Taxonomic placement of “uncertain” groups in IRMNG (for example Microsporidia in Protozoa, not Fungi; Archaeocyatha as a class of Porifera) may differ from selected other treatments	Cited totals per major group (e.g. phylum) would change if the taxa in question are allocated elsewhere
Incompletely resolved genera (not yet placed to family): approx. 103,000 names, mainly in the “uncertain” category	Will not affect currently cited totals (which are above family in this report), but listings by family will not yet be complete if generated via the IRMNG web interface.

3 Even though the true extent of this issue is not strictly knowable, an “informed guess” might put the prevalence of such names as between perhaps 1% and 5% of the names presently tallied; a “mid value” of 3% would therefore come to around 15,000 missing names. Presuming that the majority—perhaps 80%—of these are likely to be older synonyms, a very rough estimate might be that an additional 3,000 accepted, historic names might be missing from current IRMNG holdings, representing an additional approx. 1% on the present cited total of around 298,000 accepted names presently held. Allowing for an estimated 9,400 missing recently published accepted names as calculated herein, we therefore estimate that approximately 310,000 accepted generic names have been published to the end of 2019, with the total holdings of IRMNG therefore being around 96% complete in this respect.

from where their associated attributes (author, year, place of publication, etc.) can still be derived.

Desirable features of newly published lists—an IRMNG perspective

In addition to providing content to interested parties, IRMNG is of course a consumer (ingester) of content created by others in the form of published lists of names and taxonomic treatments, some of which are planned to appear in the forthcoming series of papers on the theme of “All genera of the world” in the present journal. Accordingly, IRMNG (and similar projects) will benefit considerably if the lists of taxa, etc. published in the present series can be made available as machine readable data, for example, following the Darwin Core Archive (DwCA) standard (GBIF 2017). As part of such a standard, taxonomic names and their associated authorities are disaggregated (atomised) into various Darwin Core (DwC) terms (scientificName, scientificNameAuthorship), the parent name can be indicated as both a name and an ID (parentNameUsage, parentNameUsageID), the place of publication can be indicated (namePublishedIn), and so on; for a complete list of available terms refer TDWG (2020). Both accepted and unaccepted names (synonyms and more) are equally of interest to IRMNG, and the latter can be included as separate rows in a DwCA table, with “acceptedNameUsage” and “acceptedNameUsageID” set to those for the accepted name corresponding to any name supplied as a synonym.

Additional potential interactions with the taxonomic community

To date, IRMNG content has been entered, and is being maintained, by a small number of persons with an interest in either using the data themselves, or providing it as a service to other biodiversity projects. As is the case with WoRMS, expansion of the “editor base” is always welcome and interested parties can express offers of assistance to the supplied contact point, info@irmng.org, in the first instance. The same email address can also be used for comments or other feedback on the correctness or completeness of any record. Offers to assist with the review and upgrade the taxonomic placement of names presently categorised “uncertain” are particularly welcome, and will lead to the gradual improvement of quality of IRMNG data (and summary statistics such as those presented herein) for the benefit of present and future users of the compilation. In this respect, the number of “uncertain” taxa in any group (which can be obtained by doubling the relevant plus-or-minus “range” figure given in Table 2) can give an indication of the groups most in need of additional attention in this regard.

Data availability

At time of writing, the base data (c. 492,000 accepted and unaccepted genus names plus higher taxa) from which the present totals are derived are available as a static data dump in DwCA format (file size: c. 70.5 MB, zipped) via <http://www.irmng.org/download.php>. (The same location will be used for any newer versions of the data as available; the present version will be included in the archive at <http://www.irmng.org/export/2020/>.) Instructions on use of the file once downloaded are available at the IRMNG “download” link as indicated above. The IRMNG data can also be interrogated live via the web interface, noting however that, being a dynamic dataset, some additions and alterations are to be expected over time as compared with any particular static “snapshot” (data file). Note, in this data format, quoted higher taxa are restricted to the “Linnean” ranks i.e. kingdom-*phylum*-class-order-family; for intermediate ranks such as subphylum or superclass (where implemented), relevant subsets of names can be generated via the IRMNG web interface, where such ranks are available as user selectable options.

Concluding remarks

This paper gives an indication of both the breadth and depth of coverage required for an integrated taxonomic scheme covering “all life”—also including some names that exist under nomenclatural rules but are currently considered to apply to non-biota—to the level of genus, both extant and fossil. While it is recognised that for many of these groups—in particular the extant representatives—online or print compilations exist which already catalogue genus and/or species names, such as the Eschmeyer’s Catalog of Fishes for extant fishes (Fricke *et al.* 2020), *Index Fungorum* for Fungi, etc., the compilation of data in IRMNG in a consistent format for all groups, plus its availability in a community-supported taxonomic data exchange format (Darwin Core Archive), provides a resource which supports biota-wide queries (such as “where and for what has this name previously been used”) as well as allowing the generation of lists of names according to user-specified subsetting criteria, including by taxonomic group. In addition, the existence of the IRMNG list offers a mechanism to detect homonyms, as a set of recent publications in Crustacea will attest (Low & Guinot 2010, Ng & Takeda 2010, Schweitzer *et al.* 2011, plus more), as well as providing one resource to mitigate the inadvertent creation of new homonyms.

Acknowledgements

T.R. acknowledges the generous assistance of several prior data compilers in providing content for the initial version of IRMNG, in particular Sheila Brands, the Netherlands (Systema Naturae 2000 database: Brands 1989–2005) and Ellen Farr, U.S.A. (*Index Nominum Genericorum*), as well as additional workers who have contributed content to IRMNG over the intervening 14 years. L.V, B.V. and W.D. acknowledge provision of facilities by VLIZ as well as support from LifeWatch, while facilities for the hosting of IRMNG over its initial 8 years of existence were provided by CSIRO Marine and Atmospheric Research (now CSIRO Oceans and Atmosphere) in Hobart, Tasmania, Australia. The comments of an anonymous reviewer were useful in improving the manuscript.

References

- Adl, S.M., Bass, D., Lane, C.E., Lukeš, J., Schoch, C.L., Smirnov, A., Agatha, S., Berney, C., Brown, M.W., Burki, F., Cárdenas, P., Čepička, I., Chistyakova, L., del Campo, J., Dunthorn, M., Edvardsen, B., Eglit, Y., Guillou, L., Hampl, V., Heiss, A.A., Hoppenrath, M., James, T.Y., Karnkowska, A., Karpov, S., Kim, E., Kolisko, M., Kudryavtsev, A., Lahr, D.J.G., Lara, E., Gall, L.L., Lynn, D.H., Mann, D.G., Massana, R., Mitchell, E.A.D., Morrow, C., Park, J.S., Pawlowski, J.W., Powell, M.J., Richter, D.J., Rueckert, S., Shadwick, L., Shimano, S., Spiegel, F.W., Torruella, G., Youssef, N., Zlatogursky, V. & Zhang, Q.-Q. (2019) Revisions to the classification, nomenclature, and diversity of eukaryotes. *Journal of Eukaryotic Microbiology*, 66 (1), 4–119.
<https://doi.org/10.1111/jeu.12691>
- Benton, M. (Ed.) (1993) *The Fossil Record 2*. Chapman & Hall, London, 845 pp.
- Brands, S.J. (1989–2005) *Systema Naturae 2000* (database). Amsterdam, The Netherlands. Available from: <http://sn2000.taxonomy.nl/> (copy provided to IRMNG by the author, August 2006)
- Boyle, B., Hopkins, N., Lu, Z., Raygoza Garay, J.A., Mozzherin, D., Rees, T., Matasci, N., Narro, M.L., Piel, W.H., McKay, S.J., Lowry, S., Freeland, C., Peet, R.K. & Enquist B.J. (2013) The taxonomic name resolution service: an online tool for automated standardization of plant names. *BMC Bioinformatics*, 14 (1), 16.
<https://doi.org/10.1186/1471-2105-14-16>
- Cavalier-Smith, T. (1998) A revised six-kingdom system of life. *Biological Reviews*, 73, 203–266.
<https://doi.org/10.1017/S0006323198005167>
- Cavalier-Smith, T. (2016) Higher classification and phylogeny of Euglenozoa. *European Journal of Protistology*, 56, 250–276.
<https://doi.org/10.1016/j.ejop.2016.09.003>
- Cavalier-Smith, T., Chao, E.E. & Lewis, R. (2015) Multiple origins of Heliozoa from flagellate ancestors: new cryptist subphylum Corbihelia, superclass Corbistoma, and monophyly of Haptista, Cryptista, Hacrobia and Chromista. *Molecular Phylogenetics and Evolution*, 93, 331–362.
<https://doi.org/10.1016/j.ympev.2015.07.004>
- Cavalier-Smith, T., Chao, E.E. & Lewis, R. (2016) 187-gene phylogeny of protozoan phylum Amoebozoa reveals a new class (Cutosea) of deep-branching, ultrastructurally unique, enveloped marine Lobosa and clarifies amoeba evolution. *Molecular Phylogenetics and Evolution*, 99, 275–296.
<https://doi.org/10.1016/j.ympev.2016.03.023>
- Cavalier-Smith, T., Chao, E.E. & Lewis, R. (2018) Multigene phylogeny and cell evolution of chromist infrakingdom Rhizaria: contrasting cell organisation of sister phyla Cercozoa and Retaria. *Protoplasma*, 255 (5), 1517–1574.
<https://doi.org/10.1007/s00709-018-1241-1>
- Chapman, A.D. (2009) *Numbers of Living Species in Australia and The World, 2nd edition*. Australian Biological Resources Study, Canberra, 80 pp.
- Clarivate Analytics (2020) *Index to Organism Names*. Available from: <http://www.organismnames.com/> (accessed 24 March 2020).
- Costello, M.J. (2000) Developing species information systems: the European Register of Marine Species. *Oceanography*, 13 (3), 48–55.
<https://doi.org/10.5670/oceanog.2000.09>
- Costello, M. (2015) Biodiversity: the known, unknown, and rates of extinction. *Current Biology*, 25 (9), R368–R371.
<https://doi.org/10.1016/j.cub.2015.03.051>
- Devine, A. & Coddington, J. (2019) Taxonomic gap analysis: a method of evaluating the taxa represented in biodiversity databases. *Biodiversity Information Science and Standards*, 3, e37288.
<https://doi.org/10.3897/biss.3.37288>
- Farr, E.R. & Zijlstra, G. (Eds.) (2020) *Index Nominum Genericorum (Plantarum)*. Smithsonian Institution, Washington. Available from: <https://naturalhistory2.si.edu/botany/ing/> (accessed 24 March 2020)
- Fricke, R., Eschmeyer, W.N. & van der Laan, R. (Eds.) (2020) *Eschmeyer's Catalog of Fishes: Genera, Species, References*. Available from: <http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp> (accessed 24 March 2020)
- Gawryluk, R.M.R., Tikhonenkov, D.V., Hehenberger, E., Husnik, F., Mylnikov, A.P. & Keeling, P.J. (2019) Non-photosynthetic predators are sister to red algae. *Nature*, 572, 240–243.
<https://doi.org/10.1038/s41586-019-1398-6>
- GBIF (2017) *Darwin Core Archives—How-to Guide, version 2.0*. Global Biodiversity Information Facility, Copenhagen. Available from: <https://github.com/gbif/ipt/wiki/DwCAHowToGuide> (accessed 24 March 2020)
- Guiry, M.D. & Guiry, G.M. (2020) *AlgaeBase*. World-wide electronic publication, National University of Ireland, Galway. Available from: <http://www.algaebase.org/> (accessed 24 March 2020)
- Häntzschel, W. (1975) Miscellaneous Supplement 1. Trace fossils and Problematica, second edition. In: Moore, R.C. (Ed.), *Treatise on Invertebrate Paleontology, Part W*, pp. W1–W269. Geological Society of America and University of Kansas, New York and Lawrence.
- Hauer, T. & Komárek, J. (2020) *CyanoDB 2.0 - On-line Database of Cyanobacterial Genera*. World-wide electronic publication, University of South Bohemia & Institution of Botany, Academy of Sciences of the Czech Republic. Available from: <http://www.cyanodb.cz/> (accessed 24 March 2020)
- Horton, T., Gofas, S., Kroh, A., Poore, G.C.B., Read, G., Rosenberg, G., Stöhr, S., Bailly, N., Boury-Esnault, N., Brandão, S.N., Costello, M.J., Decock, W., Dekeyser, S., Hernandez, F., Mees, J., Paulay, G., Vandepitte, L., Vanhoorne, B. & Vranken,

- S. (2017) Improving nomenclatural consistency: a decade of experience in the World Register of Marine Species. *European Journal of Taxonomy*, 389, 1–24.
https://doi.org/10.5852/ejt.2017.389
- Index Fungorum Partnership (The Royal Botanic Gardens, Kew, Landcare Research-New Zealand and the Institute of Microbiology, Chinese Academy of Science) (2020) *Index Fungorum*. Available from: <http://www.indexfungorum.org/> (accessed 24 March 2020)
- International Commission on Zoological Nomenclature (ICZN) (1999) *International Code of Zoological Nomenclature, fourth edition*. International Trust for Zoological Nomenclature, London. Available from: <https://www.iczn.org/the-code/the-international-code-of-zoological-nomenclature/the-code-online/> (accessed 24 March 2020)
- International Commission on Zoological Nomenclature (ICZN) (2012) Amendment of Articles 8, 9, 10, 21 and 78 of the International Code of Zoological Nomenclature to expand and refine methods of publication. *Zootaxa*, 3450, 1–7.
https://doi.org/10.11646/zootaxa.3450.1.1
- International Committee on Taxonomy of Viruses (ICTV) (2019) *Virus Taxonomy*. Available from: <https://talk.ictvonline.org/taxonomy/> (accessed 24 March 2020)
- International Mycological Association (2020) *Mycobank*. Available from: <http://www.mycobank.org> (accessed 24 March 2020)
- IPNI Partnership (The Royal Botanic Gardens, Kew, The Harvard University Herbaria, and The Australian National Herbarium) (2020) *International Plant Names Index (IPNI)*. Available from: <https://www.ipni.org/> (accessed 24 March 2020)
- Ivantsov, A.Y., Fedonkin, M.A., Nagovitsyn, A.L. & Zakrevskaya, M.A. (2019) *Cephalonega*, a new generic name, and the system of Vendian Proarticulata. *Paleontological Journal*, 53 (5), 447–454.
- Jain, S. (2020) Chitinozoa. In: Jain, S. *Fundamentals of Invertebrate Palaeontology*. Springer Geology, New Delhi, pp. 1–25.
https://doi.org/10.1007/978-81-322-3962-8_1
- Kansas University online Bibliography of Paleobotany (2009 version). Accessed from: <http://paleobotany.bio.ku.edu/BiblioOffPaleo.htm> (April 2011) Note: no longer accessible; last version of the home page can be viewed at <https://web.archive.org/web/20160304130504/http://paleobotany.bio.ku.edu/BiblioOffPaleo.htm>.
- Karpov, S.A., Mamkaeva, M.A., Aleoshin, V.V., Nassonova, E., Lilje, O. & Gleason, F.H. (2014) Morphology, phylogeny, and ecology of the aphelids (Aphelidea, Opisthokonta) and proposal for the new superphylum Opisthosporidia. *Frontiers in Microbiology*, 5, 112.
https://doi.org/10.3389/fmicb.2014.00112
- Kerner, A., Debrenne, F. & Vignes-Lebbe, R. (2011) Cambrian archaeocyathan metazoans: revision of morphological characters and standardization of genus descriptions to establish an online identification tool. In: Smith, V. & Penev, L. (Eds.) *e-Infrastructures for data publishing in biodiversity science*. *ZooKeys*, 150, 381–395.
https://doi.org/10.3897/zookeys.150.1566
- Knaust, D. (2020) Invertebrate coprolites and cololites revised. *Papers in Palaeontology*, 2020, 1–39.
https://doi.org/10.1002/spp2.1297
- Lax, G., Eglit, Y., Eme, L., Bertrand, E.M., Roger, A.J. & Simpson, A.G.B. (2018) Hemimastigophora is a novel supra-kingdom-level lineage of eukaryotes. *Nature*, 564, 410–414.
https://doi.org/10.1038/s41586-018-0708-8
- Low, M.E.Y. & Guinot, D. (2010) *Feldmannius* nom. nov. for *Feldmannia* Casadio, Marensi & Santillana, “30-09-2001”, preoccupied by *Feldmannia* Guinot & Tavares, “28 septembre 2001”: a case of 48-hour precedence. *Zootaxa*, 2568, 67–68.
https://doi.org/10.11646/zootaxa.2568.1.3
- Neave, S.A., and successors (Eds.) (1939–2004) *Nomenclator Zoologicus* volumes 1–9 (print) plus volume 10 (electronic). Zoological Society of London, London (print); MBLWHOI Library/uBio Project (online version). Available from: <http://ubio.org/NomenclatorZoologicus/> (accessed 24 March 2020)
- Nelson, J.S., Grande, T.C. & Wilson, M.V.H. (2016) *Fishes of the World, 5th Edition*. John Wiley & Sons, Hoboken, New Jersey, 752 pp.
- Ng, P.K. & Takeda, M. (2010) *Zehntneriana*, a replacement name for *Zehntneria* Takeda, 1972 (Crustacea, Brachyura, Pilumnidae), preoccupied by *Zehntneria* Brunner Von Wattenwyl, 1907 (Insecta, Orthoptera, Phasmidae). *Bulletin of the National Museum of Nature and Science, Series A*, 36 (2), 49–50.
- Novikoff, A. & Barabasz-Krasny, B. (2015) System of Embryophytes. In: Novikoff, A. & Barabasz-Krasny, B. *Modern Plant Systematics*. Liga-Pres, Lviv, Ukraine, pp. 23–63.
- Nozères, C., Vandepitte, L., Appeltans, W. & Kennedy, M. (2012) *Best Practice Guidelines in the Development and Maintenance of Regional Marine Species Checklists, version 1.0, released on August 2012*. Global Biodiversity Information Facility, Copenhagen, 32 pp.
- The Paleobiology Database (2020) Available from: <https://paleobiodb.org> (accessed 24 March 2020)
- Parker, C.T., Tindall, B.J. & Garrity, G.M. (Eds.) (2019) International Code of Nomenclature of Prokaryotes. Prokaryotic Code (2008 Revision). *International Journal of Systematic and Evolutionary Microbiology*, 69 (1A), S1–S111.
https://doi.org/10.1099/ijsem.0.000778
- Parker, S.P. (Ed.) (1982) *Synopsis and Classification of Living Organisms*. McGraw-Hill, New York, 2 volumes, 1166 and 1232 pp.
- Parte, A.C. (2020) *LPSN - List of Prokaryotic Names with Standing in Nomenclature*. Available from: <http://www.bacterio.net/> (accessed 24 March 2020)
- The Plant List (2013) *Version 1.1*. Available from: <http://www.theplantlist.org/> (accessed 24 March 2020)
- Raaben, M.E., Sinha, A.K. & Sharma, M. (2001) *Precambrian Stromatolites of India and Russia (A Catalogue of Type-Form-Genera)*. Birbal Sahni Institute of Palaeobotany, Lucknow, 125 pp.
- Rees, T. (2008) IRMNG—The Interim Register of Marine and Nonmarine Genera. In: Weitzman A. & Belbin, L. (Eds.) *The Proceedings of TDWG (2008), Fremantle, Australia*. Biodiversity Information Standards (TDWG) and Missouri Botanical Garden, St. Louis, p. 72.
- Rees, T. (2014) Taxamatch, an algorithm for near (‘fuzzy’) matching of scientific names in taxonomic databases. *PLoS ONE*, 9 (9), e107510.
https://doi.org/10.1371/journal.pone.0107510
- Rees, T. (Compiler) (2020) *The Interim Register of Marine and Nonmarine Genera (IRMNG)*. Available from: <http://www.irmng.org/> (accessed 24 March 2020)
- Rees, T., Vandepitte, L., Decock, W. & Vanhoorne, B. (2017) IRMNG 2006–2016: 10 years of a global taxonomic database. *Biodiversity Informatics*, 12, 1–44.
https://doi.org/10.17161/bi.v12i0.6522
- Rietschel, S. & Nitecki, M.H. (1984) Ordovician receptaculitid algae from Burma. *Paleontology*, 27, 415–420.

- Roskov, Y., Ower, G., Orrell, T., Nicolson, D., Bailly, N., Kirk, P.M., Bourgoin, T., DeWalt, R.E., Decock, W., Nieukerken, E. van, Zarucchi, J. & Penev, L. (Eds.) (2019) Species 2000 & ITIS Catalogue of Life, 2019 Annual Checklist. Species 2000: Naturalis, Leiden, the Netherlands. Available from: <http://www.catalogueoflife.org/annual-checklist/2019> (accessed 24 March 2020)
- Ruggiero, M.A. (2014) *Families of All Living Organisms, Version 2.0.a.15*, (4/26/14). Expert Solutions International, LLC, Reston, VA. 420 pp. Electronic version available from: <https://www.gbif.org/dataset/8067e0a2-a26d-4831-8a1e-21b9118a299c> (accessed 24 March 2020) <https://doi.org/10.15468/tfp6yv>
- Ruggiero, M.A., Gordon, D.P., Orrell, T.M., Bailly, N., Bourgoin, T., Brusca, R.C., Cavalier-Smith, T., Guiry, M.D. & Kirk, P.M. (2015) A higher level classification of all living organisms. *PLoS One*, 10(4), e0119248. (also correction at <https://doi.org/10.1371/journal.pone.0130114>). <https://doi.org/10.1371/journal.pone.0119248>
- Schultze-Motel, J. (2003) *Index of Generic Names of Fossil Plants, 1979–2000. Fossilium Catalogus II: Plantae pars 106*. Backhuys Publishers, Leiden, 218 pp.
- Schweitzer, C.E., Dworschak, P.C. & Martin, J.W. (2011) Replacement names for several fossil Decapoda. *Journal of Crustacean Biology*, 31 (2), 361–363. <https://doi.org/10.1651/10-3395.1>
- Sørensen, M.V., Dal Zotto, M., Rho, H.S., Herranz, M., Sánchez, N., Pardos, F. & Yamasaki, H. (2015) Phylogeny of Kinorhyncha based on morphology and two molecular loci. *PLoS One*, 10 (7), e0133440. <https://doi.org/10.1371/journal.pone.0133440>
- Taylor, T.N., Taylor, E. & Krings, M. (2009) *Paleobotany—The Biology and Evolution of Fossil Plants. Second Edition*. Elsevier/Academic Press, Amsterdam, Boston, etc., 1252 pp.
- TDWG (2020) *Darwin Core quick reference guide*. Available from: <https://dwc.tdwg.org/terms/> (accessed 24 March 2020)
- Tedersoo, L., Sánchez-Ramírez, S., Kõljalg, U., Bahram, M., Döring, M., Schigel, D., May, T., Ryberg, M. & Abarenkov, K. (2018) High-level classification of the Fungi and a tool for evolutionary ecological analyses. *Fungal Diversity*, 90 (1), 135–159. <https://doi.org/10.1007/s13225-018-0401-0>
- Turland, N.J., Wiersema, J.H., Barrie, F.R., Greuter, W., Hawksworth, D.L., Herendeen, P.S., Knapp, S., Kusber, W.-H., Li, D.-Z., Marhold, K., May, T.W., McNeill, J., Monro, A.M., Prado, J., Price, M.J. & Smith, G.F. (Eds.) (2018) *International Code of Nomenclature for Algae, Fungi, and Plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017. Regnum Vegetabile, 159*. Koeltz Botanical Books, Glashütten. <https://doi.org/10.12705/Code.2018>
- Vandepitte, L., Vanhoorne, B., Decock, W., Dekeyzer, S., Verbeeck, A.T., Bovit, L., Hernandez, F. & Mees, J. (2015) How Aphia—the platform behind several online and taxonomically oriented databases—can serve both the taxonomic community and the field of biodiversity informatics. *Journal of Marine Science and Engineering*, 3 (4), 1448–1473. <https://doi.org/10.3390/jmse3041448>
- Vandepitte, L., Vanhoorne, B., Decock, W., Vranken, S., Lanssens, T., Dekeyzer, S., Verfaillie, K., Horton, T., Kroh, A., Hernandez, F. & Mees, J. (2018) A decade of the World Register of Marine Species—general insights and experiences from the Data Management Team: where are we, what have we learned and how can we continue? *PLoS One*, 13 (4), e0194599. <https://doi.org/10.1371/journal.pone.0194599>
- Watt, A.D. (1978) *Index of Generic Names of Fossil Plants, 1974–1978 (Geological Survey Bulletin 1517)*. United States Government Printing Office, Washington, 62 pp.
- WoRMS Editorial Board (2020) *World Register of Marine Species*. Available from: <http://www.marinespecies.org> (accessed 24 March 2020). Full list of editors (Horton, T. *et al.*) available from: <http://www.marinespecies.org/aphia.php?p=popup&name=citation>. <https://doi.org/10.14284/170>
- Zhang, Z.-Q. (Ed.) (2011) Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness. *Zootaxa*, 3148 (1), 1–237. <https://doi.org/10.11646/zootaxa.3148.1>
- Zhang, Z.-Q. (Ed.) (2013) Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness (Addenda 2013). *Zootaxa*, 3703 (1), 1–82. <https://doi.org/10.11646/zootaxa.3703.1>

Supplementary file: IRMNG DwCA data file (genera and above), March 2020 version (zipped format, c. 70.5 MB), used to generate the data in this paper. For additional information refer the IRMNG “downloads” page at <https://www.irmng.org/download.php>.