

Third release of the plant rDNA database with updated content and information on telomere composition and sequenced plant genomes

Daniel Vitales^{1*}, Ugo D'Ambrosio^{1*}, Francisco Gálvez², Aleš Kovařík³, Sònia Garcia^{1**}

¹*Institut Botànic de Barcelona (IBB-CSIC- ICUB), Passeig del Migdia s/n, 08038 Barcelona, Catalonia, Spain.*

²*Bioscripts - Centro de Investigación y Desarrollo de Recursos Científicos, 41012 Sevilla, Andalusia, Spain. 4*

³*Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno CZ-61265, Czech Republic.*

*Both authors contributed equally to this work

**Corresponding author: soniagarcia@ibb.csic.es, tel. 0034 932890611.

ORCID code of the authors:

Daniel Vitales: [0000-0002-3109-5028](https://orcid.org/0000-0002-3109-5028)

Ugo D'Ambrosio: [0000-0002-5231-1762](https://orcid.org/0000-0002-5231-1762)

Francisco Gálvez: [0000-0001-5836-3881](https://orcid.org/0000-0001-5836-3881)

Aleš Kovařík: [0000-0003-2896-0698](https://orcid.org/0000-0003-2896-0698)

Sònia Garcia: [0000-0002-3143-0527](https://orcid.org/0000-0002-3143-0527)

Acknowledgements

We thank Karina Barros and Maria Pilar Alonso-Lifante, librarians of the IBB-CSIC-ICUB, for their help in publication search. The Dirección General de Investigación Científica y Técnica from the Government of Spain (CGL2016-75694-P), the Czech Science Foundation (P506/16/02149J) and the Government of Catalonia (“Ajuts a grups de recerca consolidats”, 2014SGR514) are acknowledged for funding. Sònia Garcia benefits from a “Ramón y Cajal” contract from the Government of Spain (RYC-2014-16608).

Abstract

Here we present the third release of the plant rDNA database (March 2017), an open access online resource with information on numbers, locations and structure of 5S and 18S-5.8S-26S (35S) ribosomal DNA (rDNA) (www.plantrdnadatabase.com). Data are now available for 2,148 species (3,783 entries), extracted from 785 papers published until the end of 2016. This means an expansion of 33.5% in terms of new species and 13% in new publications consulted. We appreciate an increased interest on rDNA loci research in recent years, since 10.78% of all data available were published only in 2016. The database has been expanded to include information on telomere composition and on species whose genome has been fully sequenced up to date. Telomere sequence is only known with certainty for 9.60% of species in the database and for 36.79% at the genus level, indicating, potentially, that the consensus plant telomere (*Arabidopsis*-type) might not be as extended as previously thought. We have also introduced the taxonomic category order as an additional option for data browsing. Similarly, we have included a new category to indicate the hybrid status of taxa. In addition, we upgraded and/or proofread tabs and links and slightly modified the website for a more dynamic appearance. This manuscript provides a synopsis of these changes and developments.

Keywords 5S; 35S; chromosome; karyotype; sequenced plant genomes; telomere

Running title Third release of the plant rDNA database

Introduction

The interest in chromosomes and chromosome features has been particularly conspicuous among botanists. Indeed, chromosomes were first observed by the Swiss botanist Carl Nagéli in plant cells (1842). Later on, Barbara McClintock, whose research focus was maize cytogenetics, discovered transposable elements thanks to chromosome observations (McClintock, 1955). With the advent of technical improvements, not only chromosomes could be observed and species' karyotypes described, but also chromosome markers were designed, thus allowing better characterization of karyotypes.

The first chromosome markers used for cyto-systematic and karyotyping purposes were ribosomal RNA (rRNA) genes and telomeric sequences, owing to the development of *in situ* hybridisation (Leitch et al. 1994). Both regions are fundamental components of chromosomes of all cell types, involved in ribosome production and in stabilisation of chromosome ends, respectively. Their sequence conservation, clustered arrangement and ubiquity in all life forms made them the first election as chromosome landmarks. In eukaryotes there are four rRNA genes: the 18S, 5.8S and 26S/28S, usually tandemly arranged in this order (also called 35S in plants and 45S in animals), and the 5S which is also organized in tandems and separated from the rest. However, in some organisms, cases of linked arrangement between the four genes have been described: in plants, in the sunflower family (Garcia et al. 2009, 2010) and in some gymnosperms (Galián et al. 2012; Garcia et al. 2013); in animals, in some fish and arthropods, among others (Vieira et al. 2013).

Telomeres are short repetitive sequences present at chromosome ends, where they have a protective function. Most plants are assumed to possess the so-called *Arabidopsis*-type telomere sequence (TTTAGGG) but some of them, particularly in family Asparagaceae, possess also the human-type telomere (TTAGGG) and probably different proportions of other types from animals and protozoans (Sýkorová et al. 2003). Other exceptions include the recently discovered *Cestrum*-type (TTTTTTAGG) (Peška et al. 2015) and *Allium*-type telomeres (CTCGGTTATGGG) (Fajkus et al. 2016). Telomere signals can also be detected interstitially when translocations have occurred as seen in Amaryllidaceae (Souza et al. 2016). Moreover, telomere size is variable across species and sometimes across loci (Dvořáčková et al. 2015, Fulcher et al. 2015). Therefore, telomeres can constitute a valid chromosome marker, contributing to identify karyotypes.

In 2012 we released the Plant rDNA database (www.plantrdnadatabase.com), an open access online resource giving information on the number, position and arrangement in chromosomes of rDNA loci in plants (Garcia et al. 2012), filling the gap in knowledge on the systematization of such information. Additionally, it also provided data on chromosome number, ploidy level, genome size and life cycle. The data came from published papers in plant cytogenetics, mostly based in fluorescent *in situ* hybridisation studies. In 2014, we launched a second release of the database with updated content and new features improving the user's experience with the database, such as the implementation of an automatic ideogram tool or the option of downloading results in CSV files (Garcia et al. 2014). Recently, the amount of data compiled allowed a global analysis to study relationships between rDNA locus numbers, their distribution and arrangements with chromosome numbers, genome sizes and ploidy levels (Garcia et al. 2017). One of the most relevant conclusions of the latter was that the number of rDNA

loci tends to be low, with a majority of single-locus species even in polyploid taxa. This fact was related with locus loss after polyploidy, as a part of the diploidisation process, returning polyploids to a diploid-like state over time. We also found that while 35S is located mostly terminal in chromosomes, 5S is less selective. Functionality of 35S and mobility of 5S sequences across the genome may play a role in chromosomal location of rDNAs.

The number of publications on plant cytogenetics grows increasingly faster. In a similar way, the Plant rDNA database has received attention and recent papers have cited the resource (Bedini et al. 2016; Li et al. 2016; Bolsheva et al. 2016; Sochorová et al. 2017; Volkov et al. 2017). The web page has received 10,419 visits in total and has a 27.90% of returning users (June 20th 2017), with an average number of 166 visits per month since its release. With the present article we aim to describe the updated content and to introduce some new improvements to the web page, in particular: the inclusion of information of telomere composition and data on fully sequenced plant genomes.

Database content and analysis

Since the second release (June 2013), almost four years have elapsed until the third release that we are now presenting. The present update was released in March 2017 and it contains papers published or in press until the end of 2016. Within this period, the overall database content has increased by 33.25% in number of entries, and similarly in terms of new species (33.50%) with 2,148 as total number of species available. The number of publications included has grown from 610 to 785 (28.69%). Graphs showing data and publication number increase are presented in Fig. 1. Interestingly, the year in which more data has been produced is 2016 (408 entries, or 10.78% of the whole dataset), and the last five-year period was the most productive, since almost one third of the data contained in the resource are from 2012-2016. The increase in the number of new publications and data are less pronounced than in the past release (Garcia et al. 2012) but it has to be considered that in release 2.0 many overlooked/old papers had been included. Also in the second release, we did a particular effort in collecting papers with either 5S or 35S rDNA data which had been not taken into account in release 1.0.

Once again, data belong to angiosperms (95.58%), being the group that shows the highest increase (33.28%). Eudicots represent 61.28% and monocots 37.28%, being basal eudicots the remaining 1.44% of angiosperms. We follow the taxonomic scheme of the Angiosperm Phylogeny Group (APG) IV (2016). Bryophytes have improved their representation by 6-fold, particularly thanks to a recent publication on organisation of rDNA in early land plants (Rosato et al. 2016) contributing new data for 131 taxa. The number of families has grown to 114 (35.71% increase) and genera to 540 (53.85% increase). This data can be visualized in Fig. 2.

In total, data are available for seven divisions of land plants: Bryophyta (mosses), Cycadophyta (cycads), Ginkgophyta (ginkgoes), Magnoliophyta (flowering plants), Marchantiophyta (liverworts) and Pinophyta (conifers). With the present update there are new records for 29 new families, including the well-known Araceae, Bromeliaceae, Cistaceae and Myrtaceae. Yet, some economically important plant families such as Aceraceae, Boraginaceae, Ericaceae, Hammamelidaceae, Lauraceae or Piperaceae still lack any representation in the database. On the other end of familial representation, the most common families in the database continue being the crop-rich Poaceae (18.11% of

entries), Asteraceae (11.81%), Fabaceae (10.85%), Solanaceae (5.15%), Brassicaceae (4.87%) and Amaranthaceae (3.06%) excluding duplicates (Fig. 3).

Within the database, chromosome numbers ($2n$) range from four (in *Brachycome dichromosomatica* C.R.Carter among others; Adachi et al. 1997) to 180 (only found in one species, *Agave angustifolia* Haw.; Robert et al. 2008), being the modal number $2n=24$. Ploidy levels range from one (in the gametophytic mosses and liverworts) to 20 (found in the grass *Helictochloa lusitanica* (Romero Zarco) Romero Zarco; an unresolved name which may be synonym of *Helictotrichon pratense* L., according to The Plant List <http://www.theplantlist.org/>), being the most common the diploid (63.73%), followed by the tetraploid (19.48%). Regarding the number of rDNA signals (at $2n$ level), for 35S these range between one (in *Allium fistulosum* L. among others; Hizume 1994) to 42 (in *Vella lucentina* M.B.Crespo; Rosato et al. 2015) and for 5S these range between one (in *Panicum virgatum* L. among others; Young et al. 2012) to 71 (in *Tulipa fosteriana* W.Irving; Mizouchi et al. 2007), modal values being four and six for 5S and 35S. The reconstructed ancestral land plant karyotype ($2n=16$) shows even lower values, with two 5S and 35S signals (Garcia et al. 2017).

The preferred position for 35S is the subterminal (74.84%), being interstitial in 4.58% of cases and centromeric in only 3.36%; the remaining 17.22% of records present mixed positions. Conversely, 5S has no clear predilection for a particular location, although the most common is the interstitial (34.49%) followed by the subterminal (26.46%) and the pericentromeric (20.75%), the remaining 18.3% of records presenting mixed positions. For the entries with data for both 5S and 35S, in 33.58% at least one chromosome pair has both the 5S and 35S in the same chromosome, being 72.40% of these in the same arm.

We discussed recently the significance of the preferred locations of rDNA in chromosomes (Garcia et al. 2017). Whereas in the case of the mostly subterminal 35S it was related with the functionality as nucleolar organizer regions of these genes, the more disperse 5S site location could be connected with the high mobility of 5S rRNA gene sequences, found linked to transposable elements repeatedly (Drouin and Moniz de Sá, 1995; Kapitonov and Jurka, 2003; Kalendar et al. 2008; Wicke et al. 2011) or as extrachromosomal circles (Cohen et al. 2010).

With respect to the arrangement of rRNA genes in chromosomes, most taxa (90.69% of entries) present the unlinked arrangement, in which 5S and 35S rRNA genes are physically separated while only 3.07% show the linked arrangement. In the remainder 6.24% of entries, images from in situ hybridization suggest that both rDNAs are located in close proximity and may be linked, although physical linkage remains to be proven at the molecular level. As shown by Garcia et al. 2017, the separate arrangement of rRNA genes was reconstructed as the ancestral condition in land plants. Its dominant prevalence, not only in plants but also in animals, may be related to the transcription of every rDNA occurring in different cellular compartments and for which a separate organisation would be beneficial (Layat et al. 2012).

New data included: telomere composition and sequenced plant genomes

The database has newly incorporated information on which plant genomes have been fully sequenced (until February 2017) and a new tab has been created organizing these

species by families. With the advent of NGS technology, more plant genomes are being sequenced and this information is made easily accessible to users directly from the webpage. Within the catalogue, each species name has a direct hyperlink to the source publication of the corresponding genome. Out of the species present in the database, the genome has been fully sequenced for 88 of them, representing 468 entries (12.37% of the database). Certainly, most of the species with a full genome sequenced are model plants or crops from Brassicaceae, Fabaceae, Poaceae and Solanaceae. These are also the most represented in the plant rDNA database, with the significant exception of Asteraceae with only two species fully sequenced to date (*Conyza canadensis* (L.) Cronquist, synonym of *Erigeron canadensis* L., and *Cynara cardunculus* L.) besides being one of the most crop-rich and species-rich plant families.

In addition to information to fully sequenced genomes, the database also includes, as novelty, evidence on telomere composition. Data on telomere motives has been extracted from the same source publication of rDNA-FISH, in the cases where this information was also included. In addition, other publications have been consulted, addressing telomere composition in relatively large datasets representing different plant groups (i.e. Thomas et al. 1996; Adams et al. 2001; Sýkorová et al. 2003a, 2003b; Suzuki, 2004; Shibata and Hizume, 2011). Different telomere sequences have been considered, namely: *Arabidopsis*-type, *Allium*-type, *Cestrum*-type, *Bombyx*-type, *Human*-type, *Oxythricha*-type and *Tetrahymena*-type. The four latter (from animals and protozoa) are included because hybridisation signals of these sequences have been found in some plant species, occurring together with the *Arabidopsis*-type in different proportions (Sýkorová et al. 2003a), although these might correspond in some cases to cross-hybridisation artefacts. The absence of *Arabidopsis*-type is also occasionally reported, such as in species from the genera *Sessea*, *Vestia* (Solanaceae) and *Oenothera* (Onagraceae) in which the *Arabidopsis*-type telomere probe did not hybridize with the chromosome spreads (Peška et al. 2008; Golzyck et al. 2014).

In order to organize the data on telomere composition, some assumptions have been made. When a given telomere sequence has been confirmed for several species of a genus, it is considered that any other species from the same genus would present the same telomeres, although we indicate that this is an assumption based on another species. However, very exceptionally, some genera present species with different telomeric repeats such as the carnivorous *Genlisea* (Tran et al. 2015). In the same line, we state that a given species may have “probably” or “possibly” a given telomere type when there is evidence at the family or order level, respectively, indicating in which species the assumption is based on. Finally, when there is no evidence above the taxonomic level order we state that the telomere composition is unknown (except in bryophytes and pteridophytes, for which we consider the level division instead of the order category).

Considering this, the telomere composition is confidently known for 16.90% of the entries, representing only 9.60% of the species in the database. However, if we accept that telomeres rarely change at the genus level, 64.77% of the entries could be considered as known regarding telomere composition, corresponding to 36.79% of the species. In other words, 63.21% of the species in the database have fairly unknown telomere composition. Out of the aforementioned species with known telomeres, 24.59% present *Arabidopsis*-type, 11.27% present a mixture of telomere sequences (mostly human and *Arabidopsis*), 0.65% are *Allium*-type and 0.28% are *Cestrum*-type (in terms of database entries: 43.29%, 19.84%, 1.14% and 0.50% respectively). In view

of these results, the telomere field of research could potentially expand to uncover new plant telomere sequences in the coming years, and the database could be used as a starting point to select candidate species for further research.

Complementary information

In order to complement the primary information (chromosome number, ploidy level and numbers and positions of 5S and 35S rDNA loci) the database provides additional data. Genome sizes (2C) are supplied for 48.61% of entries (Garnatje et al. 2011 and Bennett and Leitch 2012), ranging between 0.13 (*Genlisea margaretae* Hutch.) and 125.67 pg (*Fritillaria imperialis* L.). Information on life cycle is also included when available (69.98% of entries in the database), being the perennial the most common (63.13%). The presence of secondary constrictions, an indication of NOR activity, is described in some papers reporting rDNA-FISH. The database reports this information for 50.01% of the entries included.

Other improvements to the database

One of the new possibilities that the database offers is the inclusion of the order taxonomic category. In this regard, data are now available for 53 orders, and data can be browsed by this category too. New orders included in this update are mostly from bryophytes (Dicranales, Funariales, Hypnales, Jungermanniales, Lunulariales, Polytrichiales, Porellales and Pottiales) and the eudicot order Oxalidales. Following the recent edition of the APG IV system (2016), family Xanthorreaceae has been changed to Asphodelaceae. Also, species from orders Ranunculales and Proteales have been placed as basal eudicots whereas the remaining eudicots have been renamed as core eudicots.

In the previous editions of the plant rDNA database, the hybrid status of a given entry was indicated in the field “observations” together with other diverse information that could be interesting (e.g. name or locality of the population studied, accession number, any cytogenetic particularity like presence of B-chromosomes, etc.). However, we thought that it could be interesting that information on hybrid origin was available as an option in the database due to the fact that hybridization processes can impact rDNA loci, both in numbers and distribution in chromosomes. Therefore, in this new version of the database, this information has been extracted and placed independently, and eventually further analysis could be performed. Additionally, hybrids are classified in different types: intervarietal, interspecific and intergeneric. The latter is particularly relevant for certain crops such as *Triticale* (*Triticum* × *Secale*) or *Festulolium* (*Festuca* × *Lolium*).

Besides, the web page has been thoroughly reviewed and mistakes have been corrected. Broken hyperlinks have been replaced and outdated pages in the “links” tab have been deleted. Other webpages, such as the recent Chromosome Counts Database (<http://ccdb.tau.ac.il/home/>), a new resource providing updated chromosome counts for plant species (Rice et al. 2015), have been added to the list of interesting resources.

Conclusion and future directions

The third release of the Plant rDNA database witnesses the still rising interest among the plant cytogenetics community on data on rDNA loci number, distribution and structure, despite being one of the first chromosomal markers to be used (the first paper on plant rDNA FISH dates back to 1974). As far as we are aware, the database is the most comprehensive resource available for information on 5S and 35S rDNA loci in plants, and the only one providing details on telomere composition, which may help guiding future research in order to find possible new telomere motives. Besides, a parallel initiative from our research team presenting the same kind of information for animals (www.animalrdnadatabase.com, J. Sochorová et al. under review) will soon be released. Further work will establish comparisons between trends found in plants and animals and may improve our understanding of the evolution of these vital DNA regions. Last but not least, the inclusion of which plant genomes have been fully sequenced up to date is also a helpful resource, not only for researchers interested in rDNA loci (which will know if their plant of interest has been fully sequenced or not) but as an updated and user-friendly summary tracking published plant genomes for the plant research community. The significant increase in available data on rDNA loci observed, the boost in knowledge about plant telomere sequences and the likely expansion of new plant species whose genome will be sequenced in the coming years, trigger our determination to perform regular content updates to our resource, which now provides more complete information on plant karyotypes and genomic composition.

The authors declare no conflict of interest.

References

- Adachi J, Watanabe K, Fukui K, Ohmido N, Kosuge K (1997) Chromosomal location and reorganization of the 45S and 5S rDNA in the *Brachyscome lineariloba* complex (Asteraceae). J. Pl. Res. 110: 371–377. doi: [10.1007/BF02524936](https://doi.org/10.1007/BF02524936)
- Adams SP, Hartman TPV, Lim KY, Chase MW, Bennett MD, Leitch IJ, Leitch AR (2001) Loss and recovery of *Arabidopsis*-type telomere repeat sequences 5'-(TTTAGGG)_n-3' in the evolution of a major radiation of flowering plants. Proc Roy Soc B-Biol Sci 268:1541–1546. doi: [10.1098/rspb.2001.1726](https://doi.org/10.1098/rspb.2001.1726)
- Bedini G, Pierini B, Roma-Marzio F, Caparelli KF, Bonari G, Dolci D, Gestri G, D'Antraccoli M, Peruzzi L (2016) Wikiplantbase #Toscana, breaking the dormancy of floristic data. Plant Biosyst 150:601–610. doi: [10.1080/11263504.2015.1057266](https://doi.org/10.1080/11263504.2015.1057266)
- Bennett MD, Leitch IJ (2012) Plant DNA C-values database (release 6.0, Dec. 2012). Available at: <http://www.kew.org/cvalues/>, accessed January 30th 2017
- Bolsheva NL, Dyachenko OV, Samatadze TE, Rachinskaya OA, Zakharchenko NS, Shevchuk TV, Amosova AV, Muravenko O, Zelenin AV (2016) Karyotype of *Mesembryanthemum crystallinum* (Aizoaceae) studied by chromosome banding, FISH with rDNA probes and immunofluorescence detection of DNA methylation. Plant Biosyst 150:916–922. doi: [10.1080/11263504.2014.991360](https://doi.org/10.1080/11263504.2014.991360)
- Castorena-Sánchez I, Escobedo RM, Quiroz A (1991) New cytotaxonomical determinants recognized in six taxa of *Agave* in the sections *Rigidae* and *Sisalanae*. Canad. J. Bot. 69:1257–1264. doi: [10.1139/b91-163](https://doi.org/10.1139/b91-163)
- Cohen S, Agmon N, Sobol O, Segal D (2010) Extrachromosomal circles of satellite repeats and 5S ribosomal DNA in human cells. Mobile DNA 1:11. doi: [10.1186/1759-8753-1-11](https://doi.org/10.1186/1759-8753-1-11)
- Drouin G, De Sa MM (1995) The concerted evolution of 5S ribosomal genes linked to the repeat units of other multigene families. Mol Biol Evol 12:481–493. doi: [10.1186/1471-2148-11-151](https://doi.org/10.1186/1471-2148-11-151)
- Dvořáčková M, Fotjová M and Fajkus J (2015) Chromatin dynamics of plant telomeres and ribosomal genes. The Plant Journal 83:18–37. doi: [10.1111/tpj.12822](https://doi.org/10.1111/tpj.12822)
- Fajkus P, Peška V, Sitová Z, Fulnečková J, Dvořáčková M, Gogela R, Fajkus J (2016) *Allium* telomeres unmasked: the unusual telomeric sequence (CTCGGTTATGGG)_n is synthesized by telomerase. Plant J 85:337–347. doi: [10.1111/tpj.13115](https://doi.org/10.1111/tpj.13115)
- Fulcher N., Teubenbacher A, Kerdaffrec E, Farlow A, Nordborg M, Riha K (2015) Genetic Architecture of Natural Variation of Telomere Length in *Arabidopsis thaliana*. Genetics 199(2):625–635. doi: [10.1534/genetics.114.172163](https://doi.org/10.1534/genetics.114.172163)
- Galián JA, Rosato M, Rosselló JA (2012) Early evolutionary colocalization of the nuclear ribosomal 5S and 45S gene families in seed plants: evidence from the living fossil gymnosperm *Ginkgo biloba*. Heredity 108:640–646. doi: [10.1038/hdy.2012.2](https://doi.org/10.1038/hdy.2012.2)

- Garcia S, Kovařík A, Leitch AR, Garnatje T (2017) Cytogenetic features of rRNA genes across land plants: analysis of the Plant rDNA database. *Plant J* 89:1020–1030. doi: [10.1111/tpj.13442](https://doi.org/10.1111/tpj.13442)
- Garcia S, Gálvez F, Gras A, Kovařík A, Garnatje T (2014) Plant rDNA database: update and new features. *Database* 2014:bau063. doi: [10.1093/database/bau063](https://doi.org/10.1093/database/bau063)
- Garcia S, Kovařík A (2013) Dancing together and separate again: gymnosperms exhibit frequent changes of fundamental 5S and 35S rRNA gene rDNA organisation. *Heredity* 111:23–33. doi: [10.1038/hdy.2013.11](https://doi.org/10.1038/hdy.2013.11)
- Garcia S, Lim KY, Chester M, Garnatje T, Pellicer J, Vallès J, Leitch AR, Kovařík A (2009) Linkage of 35S and 5S rRNA genes in *Artemisia* (family Asteraceae): first evidence from angiosperms. *Chromosoma* 118:85–97. doi: [10.1007/s00412-008-0179-z](https://doi.org/10.1007/s00412-008-0179-z)
- Garcia S, Panero JL, Siroky J, Kovařík A (2010) Repeated reunions and splits feature the highly dynamic evolution of 5S and 35S ribosomal RNA genes rDNA in the Asteraceae family. *BMC Plant Biol* 10: 176. doi: [10.1186/1471-2229-10-176](https://doi.org/10.1186/1471-2229-10-176)
- Garnatje T, Canela MA, Garcia S, Hidalgo O, Pellicer J, Sánchez-Jiménez I, Siljak-Yakovlev S, Vitales D, Vallès J (2011) GSAD: A genome size in the Asteraceae database. *Cytometry Part A*. doi: [10.1002/cyto.a.21056](https://doi.org/10.1002/cyto.a.21056)
- Kalendar R, Tanskanen J, Chang W, Antonius K, Sela H, Peleg O, Schulman AH (2008) Cassandra retrotransposons carry independently transcribed 5S RNA. *Proc Natl Acad Sci USA* 105:5833–5838. doi: [10.1073/pnas.0709698105](https://doi.org/10.1073/pnas.0709698105)
- Kapitonov VV, Jurka J (2003) A novel class of SINE elements derived from 5S rRNA. *Mol Biol Evol* 20:694–702. doi: [10.1093/molbev/msg075](https://doi.org/10.1093/molbev/msg075)
- Layat E, Sáez-Vásquez J, Tourmente S (2012) Regulation of Pol I-transcribed 45S rDNA and Pol III-transcribed 5S rDNA in *Arabidopsis*. *Plant Cell Physiol* 53:267–276. doi: [10.1093/pcp/pcr177](https://doi.org/10.1093/pcp/pcr177)
- Leitch AR, Schwarzacher T, Jackson D, Leitch IJ (1994) *In situ* hybridization: a practical guide. Oxford: Bios Scientific Publishers Ltd.
- Li KP, Wu YX, Zhao H, Wang Y, Lü XM, Wang JM, Xu Y, Li ZY, Han YH (2016) Cytogenetic relationships among *Citrullus* species in comparison with some genera of the tribe Benincaseae (Cucurbitaceae) as inferred from rDNA distribution patterns. *BMC Evol Biol* 16:85. doi: [10.1093/pcp/pcr177](https://doi.org/10.1093/pcp/pcr177)
- McClintock B (1955) Controlled mutation in maize. *Carnegie Inst. Washington Year Book* 54: 245-255.
- Peška V, Fajkus P, Fojtová M, Dvořáčková M, Hapala J, Dvořáček V, Fajkus J (2015) Characterisation of an unusual telomere motif (TTTTTTAGGG)_n in the plant *Cestrum elegans* (Solanaceae), a species with a large genome. *Plant J* 82:644–654. doi: [10.1111/tpj.12839](https://doi.org/10.1111/tpj.12839)
- Peška V, Sýkorová E, Fajkus J (2008) Two faces of Solanaceae telomeres: a comparison between *Nicotiana* and *Cestrum* telomeres and telomere-binding proteins. *Cytog Gen Res* 122:380–387. doi: [10.1159/000167826](https://doi.org/10.1159/000167826)

- Rice A, Glick L, Abadi S, Einhorn M, Kopelman NM, Salman-Minkov A, Mayzel J, Chay O, Mayrose, I (2015) The Chromosome Counts Database (CCDB) – a community resource of plant chromosome numbers. *New Phytol* 206:19–26. doi: [10.1111/nph.13191](https://doi.org/10.1111/nph.13191)
- Rosato M, Kovařík A, Garilleti R, Rosselló JA (2016) Conserved organisation of 45S rDNA sites and rDNA gene copy number among major clades of early Land Plants. *PLoS ONE* 11:e0162544. doi: [10.1371/journal.pone.0162544](https://doi.org/10.1371/journal.pone.0162544)
- Shibata F, Hizume M (2011) Survey of *Arabidopsis*- and Human-type telomere repeats in plants using fluorescence in situ hybridisation. *Cytologia* 76:353–360. doi: doi.org/10.1508/cytologia.76.353
- Sochorová J, Coriton O, Kuderová A, Lunerová J, Chèvre AM, Kovařík A (2017) Gene conversion events and variable degree of homogenization of rDNA loci in cultivars of *Brassica napus*. *Ann Bot* 116:13–26. doi: [10.1093/aob/mcw187](https://doi.org/10.1093/aob/mcw187)
- Souza G, Vanzela AL, Crosa O, Guerra M (2016) Interstitial telomeric sites and Robertsonian translocations in species of *Ipheion* and *Nothoscordum* (Amaryllidaceae). *Genetica* 144(2):157–66. doi: [10.1007/s10709-016-9886-1](https://doi.org/10.1007/s10709-016-9886-1)
- Suzuki K (2004) Characterization of telomere DNA among five species of pteridophytes and bryophytes. *J Bryol* 26:175–180. doi: [10.1179/037366804X5279](https://doi.org/10.1179/037366804X5279)
- Sýkorová E, Lim KY, Chase MW, Knapp S, Leitch IJ, Leitch AR, Fajkus J (2003) The absence of *Arabidopsis*-type telomeres in *Cestrum* and closely related genera *Vestia* and *Sessea* (Solanaceae): first evidence from eudicots. *Plant J* 34:283–291. doi: [10.1046/j.1365-3113X.2003.01731.x](https://doi.org/10.1046/j.1365-3113X.2003.01731.x)
- Sýkorová E, Lim KY, Kunická Z, Chase MW, Bennett MD, Fajkus J, Leitch AR (2003) Telomere variability in the monocotyledonous plant order Asparagales. *P Roy Soc B-Biol Sci* 270:1893–1904. doi: [10.1098/rspb.2003.2446](https://doi.org/10.1098/rspb.2003.2446)
- The Angiosperm Phylogeny Group (2016) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot J Linn Soc* 181:1–20. doi: [10.1111/boj.12385](https://doi.org/10.1111/boj.12385)
- The Plant List (2010). Version 1. Published on the Internet. Available at: <http://www.theplantlist.org/>, accessed January 30th 2017
- Thomas HM, Williams K, Harper JA (1996) Labelling telomeres of cereals, grasses and clover by primed in situ DNA labelling. *Chrom Res* 4:182–184. doi: [10.1007/BF02254956](https://doi.org/10.1007/BF02254956)
- Tran TD, Cao HX, Jovtchev G, Neumann P, Novák P, Fojtová M, Vu GTH, Macas J, Fajkus J, Schubert I, Fuchs J (2015) Centromere and telomere sequence alterations reflect the rapid genome evolution within the carnivorous plant genus *Genlisea*. *Plant J* 84:1087–1099. doi: [10.1111/tpj.13058](https://doi.org/10.1111/tpj.13058)
- Vierna J, Wehner S, Honer zu Siederdisen CH, Martínez-Lage A, Marz M (2013) Systematic analysis and evolution of 5S ribosomal DNA in metazoans. *Heredity* 1115:410–421. doi: [10.1038/hdy.2013.63](https://doi.org/10.1038/hdy.2013.63)

Volkov RA, Panchuk II, Borisjuk NV, Hosiawa-Baranska M, Maluszynska J, Hemleben V (2017) Evolutional dynamics of 45S and 5S ribosomal DNA in ancient allohexaploid *Atropa belladonna*. BMC Pl Biol 17:21. doi: [10.1186/s12870-017-0978-6](https://doi.org/10.1186/s12870-017-0978-6)

Wicke S, Costa A, Muñoz J, Quandt D (2011) Restless 5S: the re-arrangement s and evolution of the nuclear ribosomal DNA in land plants. Mol Phylogenet Evol 61:321–332. doi: [10.1016/j.ympev.2011.06.023](https://doi.org/10.1016/j.ympev.2011.06.023)

Fig. captions

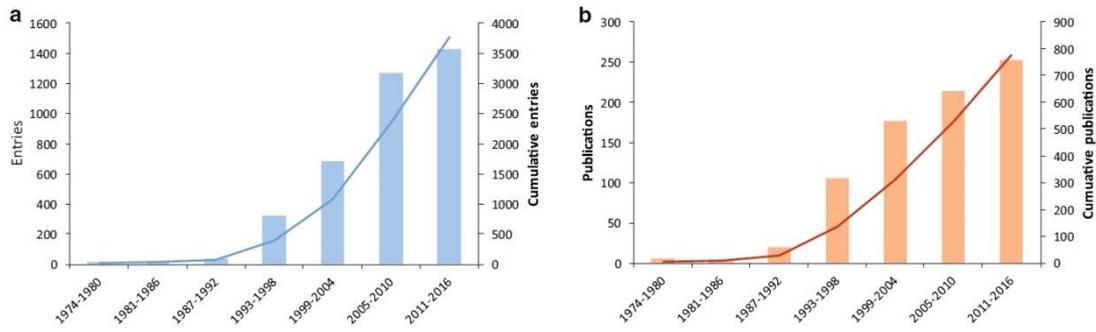


Fig. 1. Number of entries (a) and publications (b) on rDNA-FISH reported per year over seven successive 6-year periods between 1974 and 2016, the first period comprising seven years.

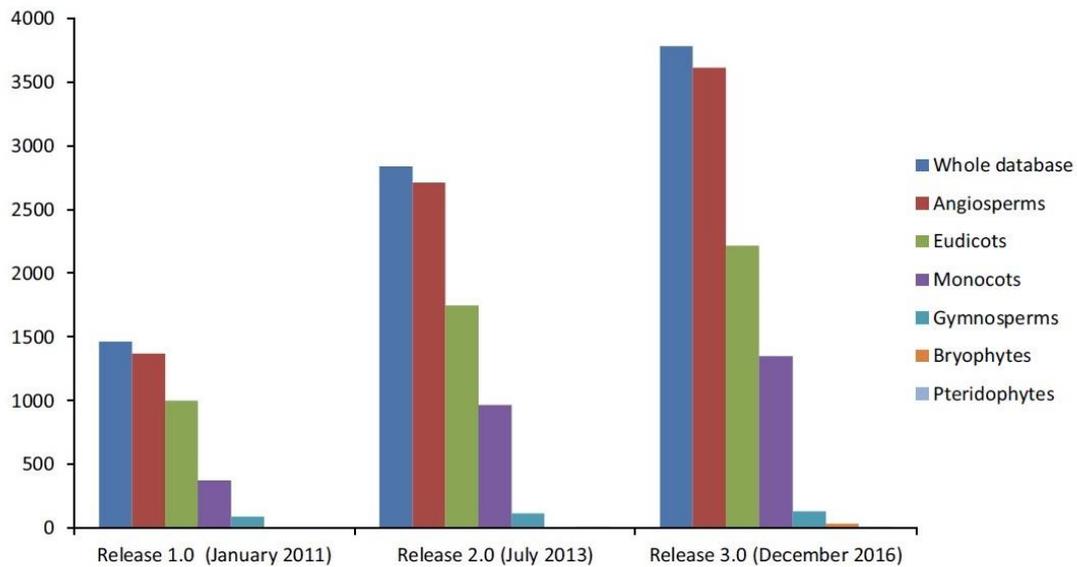


Fig. 2. Growth of the database content in terms of the number of entries in the three successive releases.

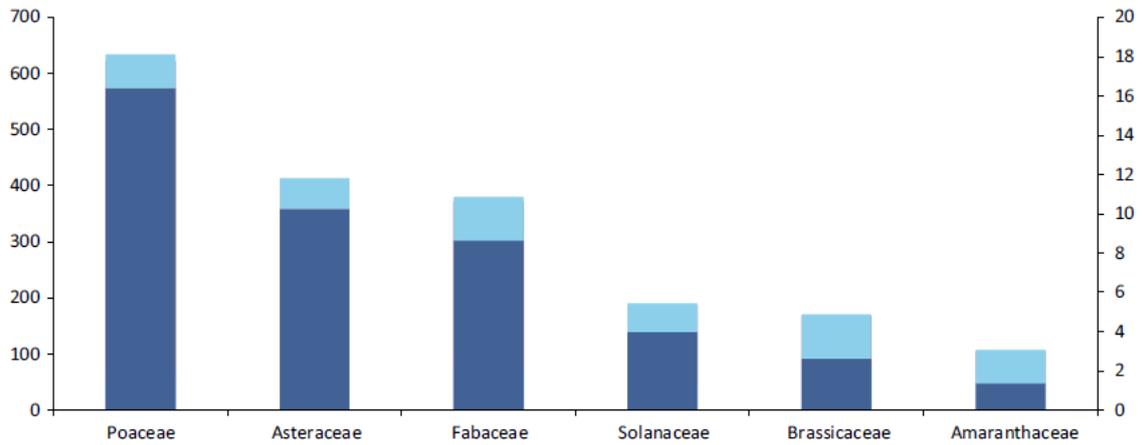


Fig. 3. The most represented families in the database in terms of number of entries (left axis) and percentage with respect to the whole database (right axis), highlighting in light blue the proportion of records of the most popular genus in each family, excluding duplicated data: *Triticum* for Poaceae (9.50% of entries), *Artemisia* for Asteraceae (13.30%), *Arachis* for Fabaceae (20.43%), *Nicotiana* for Solanaceae (26.34%), *Brassica* for Brassicaceae (46.11%), *Chenopodium* for Amaranthaceae (55.23%).