scAPAdb: a comprehensive database of alternative polyadenylation at single-cell resolution

Sheng Zhu[1,2], Qiwei Lian[2], Wenbin Ye[2], Wei Qin[2], Zhe Wu[2], Guoli Ji[2] and Xiaohui Wu [ORCID][1,*]

[1]Pasteurien College, Soochow University, Suzhou, Jiangsu 215000, China and [2]Department of Automation, Xiamen University, Xiamen, Fujian 361005, China

ABSTRACT

Alternative polyadenylation (APA) is a widespread regulatory mechanism of transcript diversification in eukaryotes, which is increasingly recognized as an important layer for eukaryotic gene expression. Recent studies based on single-cell RNA-seq (scRNA-seq) have revealed cell-to-cell heterogeneity in APA usage and APA dynamics across different cell types in various tissues, biological processes and diseases. However, currently available APA databases were all collected from bulk 3′-seq and/or RNA-seq data, and no existing database has provided APA information at single-cell resolution. Here, we present a user-friendly database called scAPAdb (http://www.bmibig.cn/scAPAdb), which provides a comprehensive and manually curated atlas of poly(A) sites, APA events and poly(A) signals at the single-cell level. Currently, scAPAdb collects APA information from > 360 scRNA-seq experiments, covering six species including human, mouse and several other plant species. scAPAdb also provides batch download of data, and users can query the database through a variety of keywords such as gene identifier, gene function and accession number. scAPAdb would be a valuable and extendable resource for the study of cell-to-cell heterogeneity in APA isoform usages and APA-mediated gene regulation at the single-cell level under diverse cell types, tissues and species.

INTRODUCTION

mRNA alternative polyadenylation (APA) is a critical mechanism that generates distinct 3′ UTR (Untranslated Region) isoforms for the same genes, which plays a significant role in influencing mRNA stability, translation efficiency and cellular localization (1,2). Nearly 70% of eukaryotic genes possess multiple polyadenylation [poly(A)] sites, offering the possibility of regulating gene expression through APA. 3′ UTR shortening/lengthening mediated by APA is involved in a range of biological processes and diseases, such as cell proliferation, differentiation, cancer and neurodegenerative disorders, through the binding of RNA-binding proteins or the escape of elements like miRNA-binding sites or AU/GU-rich elements in alternative 3′ UTRs (2,3). A number of studies based on direct 3′ end sequencing (3′-seq) or bulk RNA-seq have documented the global landscape of APA in various species and revealed that APA is dynamically modulated in a tissue and/or cell-type specific manner (reviewed in (2,4,5)).

The rapid rise of single-cell RNA sequencing (scRNA-seq) has provided a powerful means for characterizing transcriptome of individual cells at unprecedented throughput and resolution (6). The single-cell landscape of APA in various tissues, biological processes and diseases has been extensively explored, such as in acute myeloid leukaemia (7), neuronal differentiation and autism (8), secretory cell differentiation (9) and developing mouse embryo (10). Recently, computational approaches including APA-seq (11), scAPA (12), Sierra (13), scAPAtrap (14) and scDaPars (15) have been proposed to identify poly(A) sites and/or APA events at the single-cell level from standard scRNA-seq protocols, such as CEL-seq (16), Drop-seq (17) and 10× Genomics (18). A large number of APA sites in different cell types were captured from scRNA-seq, a considerable part of which were undetectable in previous bulk RNA-seq or 3′-seq data. The compendium of APA sites in single cells would bring the promise of investigating both common and rare cell types and contribute to the understanding of dynamic gene expression regulation at single-cell and isoform resolution.

Currently, several poly(A) site databases are available, all of which were collected from bulk 3′-seq and/or RNA-seq data. Databases such as APADB (19), APASdb (20), PolyA_DB 3 (21), PlantAPAdb (22) and PolyAsite 2.0 (23) were built upon 3′-seq data, data of which are generally of high quality but only covered a small number of species, tissues and/or disease samples. Additional databases, such as TC3A (24), Animal-APAdb (25), TREND-DB (26) and APAatlas (27), collected poly(A) sites and/or APA events from bulk RNA-seq by using algorithms like QAPA (28)

*To whom correspondence should be addressed. Tel: +86 13459029440; Email: xhwu@suda.edu.cn

and DaPars (29). These databases have provided a valuable repertory of APA events in plants, animals or tumors, however, no existing database provides comprehensive APA information at single-cell resolution.

The great advance in scRNA-seq has led to fast accumulation of massive amounts of publicly available sequencing data, which provides great potential to profile APA dynamics at the single-cell level. However, these large amount of deposited data in public archives not only require significant pre-processing (i.e. quality control, data curation and cell type annotation), but also need the use of a variety of relatively complex tools to extract and visualize APA information. There is a strong need for efforts involving easy access to single-cell APA information from various scRNA-seq experiments. Here we present a user-friendly database called scAPAdb (http://www.bmibig.cn/scAPAdb), which provides a comprehensive and manually curated atlas of poly(A) sites, APA events and poly(A) signals at single-cell resolution. Currently, scAPAdb records APA information in >1 million cells from 363 experiments based on a large volume of scRNA-seq data and covers six species including *Homo sapiens* (human) and *Mus musculus* (mouse), *Oryza sativa* L. (rice) Japonica and Indica, *Arabidopsis thaliana*, *Zea mays* (corn) and *Chlamydomonas reinhardtii* (Chlamy). scAPAdb details poly(A) sites and APA events with annotation information including the respective genes, coordinates and genomic regions, as well as APA information in individual cells including number of supported reads, relative APA usage and weighted 3′ UTR length. Moreover, the single-cell profile of gene expression, poly(A) signals and related sequences for poly(A) sites across cell types in different species are also available. scAPAdb also provides batch download of data, and users can query the database by gene identifier, gene symbol/alias, gene function, gene ontology and accession number. In particular, scAPAdb is equipped with a wealth of back-end programs and standard pipelines, which makes it highly scalable to integrate emerging scRNA-seq data in the future. As a user-friendly database, scAPAdb would be a valuable and extendable resource for the study of cell-to-cell heterogeneity in APA isoform usages and APA-mediated gene regulation at the single-cell level under diverse cell types, tissues and species.

DATA COLLECTION AND PROCESSING

We collected published 3′ tag-based scRNA-seq data mainly by searching PubMed (https://pubmed.ncbi.nlm.nih.gov/), bioRxiv (https://www.biorxiv.org/), PanglaoDB (30) and the 10x Genomics website (Figure 1). For datasets without cell type annotations from the original study or the PanglaoDB, we inferred cell types from the gene-cell expression matrix. We used Sierra (13) and scAPAtrap (14) for identifying poly(A) sites from scRNA-seq data. Poly(A) sites were then annotated using the movAPA package (31). A poly(A) site expression matrix (hereinafter referred as *PA-matrix*) was obtained for each experiment, with each row being a poly(A) site and each column being a cell (Figure 1). Poly(A) sites expressed in less than three cells were discarded and cells with less than 200 expressed poly(A) sites were also removed. Based on the *PA-matrix*, we quantified the APA usage for APA genes in individual cells with two APA metrics, the percentage of the proximal poly(A) site us-

age index (PPUI) and the weighted 3′ UTR length (WUL) (Figure 1). A smaller PPUI or a higher WUL score of an APA gene indicates gain of the distal poly(A) site usage, i.e. 3′ UTR lengthening of the gene. Motifs surrounding poly(A) sites were identified by DREME (32) for individual cell types in each study of each species.

In scAPAdb, we catalogue APA datasets by experiments and studies, respectively. An experiment is a dataset from cells originating from the same common biological source or sequencing experiment. A study is a dataset containing multiple experiments. To compile an APA atlas for a study, poly(A) sites from individual experiments were first annotated with known poly(A) sites collected from bulk 3′-seq to construct a reference poly(A) site dataset for each species. Then poly(A) sites for individual studies of were obtained based on the same reference. More details about data collection and processing are described in Supplementary Material. Detailed commands for data processing were provided in the 'Help' page of our scAPAdb website.

WEB INTERFACE

scAPAdb provides a user-friendly web interface with nine modules, including 'Search', 'Experiment', 'Study', 'PolyA Site', 'APA Usage', '3′ UTR Length', 'PolyA Signal', 'Statistics' and 'Download' (Figure 1). In the 'Search' module, users can query a gene or an experiment/study with various kinds of keywords to view full APA information of a gene, an experiment or a study. The 'Experiment' page catalogues the APA atlas by experiment, where each experiment is displayed as a data card. The 'Study' page catalogues the APA atlas by study. Users can filter an experiment or a study of interest by different entries such as species, sequencing protocols, tissues and release year. The 'PolyA Site', 'APA Usage' and '3′ UTR Length' modules provide APA information in individual cells of each experiment, including poly(A) sites identified by scAPAtrap and Sierra, PPUI scores and the weighted 3′ UTR length of APA genes. The 'PolyA signal' module provides sequences and comprehensive poly(A) signal motifs for different cell types in different species. The 'Download' page provides batch download of APA data for individual experiments and studies. The 'Statistics' page provides the summary of the database, as well as the descriptive information of each experiment and study.

DATABASE CONTENT AND USAGE

Atlas of single-cell APA from six species

Currently, scAPAdb provides APA datasets from 363 experiments of 86 studies in six species; most datasets were collected from mouse, human and Arabidopsis (Table 1). The 'Statistics' page of scAPAdb summarizes curated poly(A) site datasets and presents information about numbers of studies, experiments, tissues, cell types and cells per species. Full information of each experiment and study, including species, accession number, sample size, tissue, numbers of poly(A) sites and APA events, is also provided in the 'Statistics' page and in Supplementary Tables S1 and S2. Users can download the APA data of each experiment or study in batches from the 'Download' page.
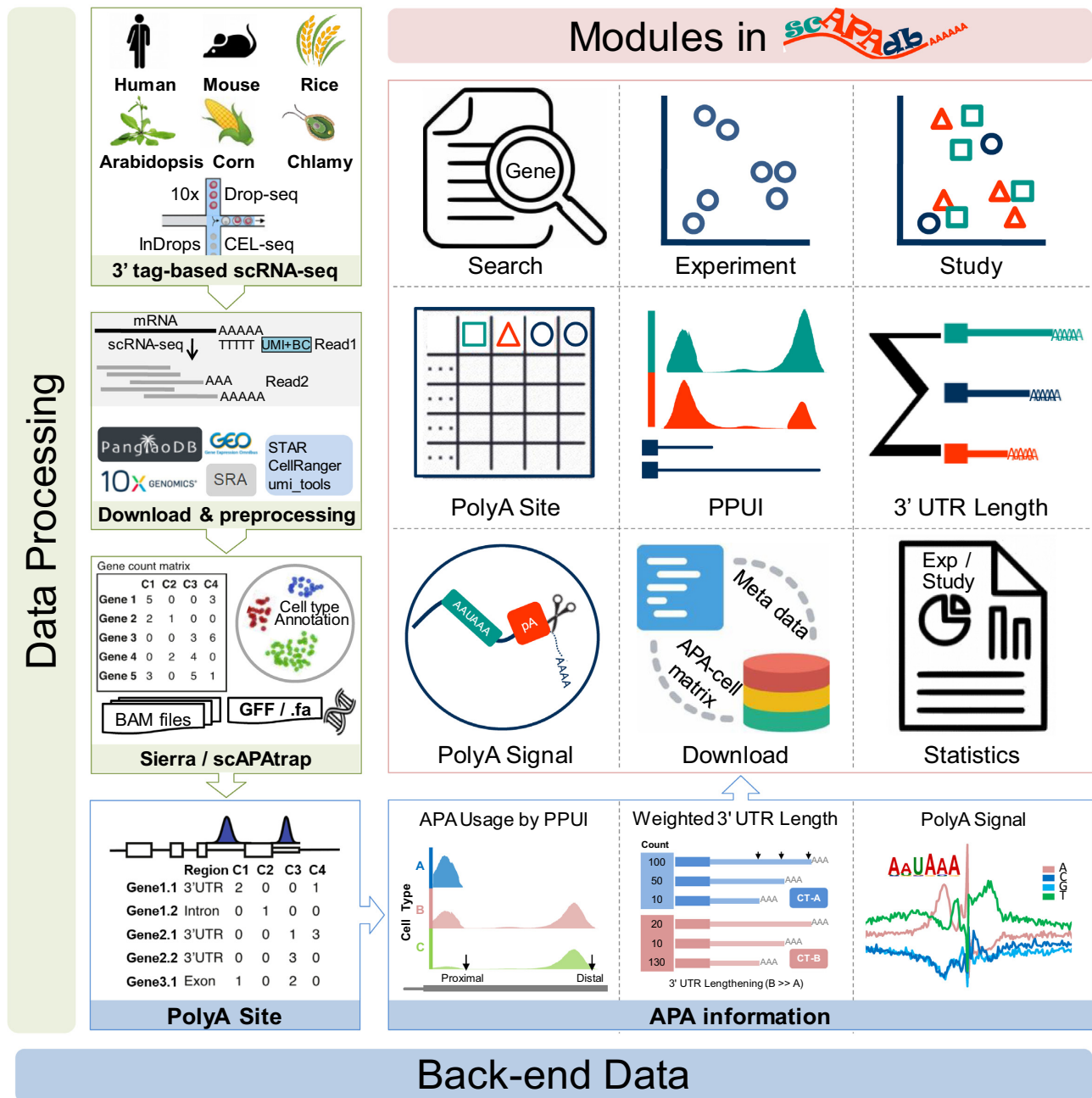
Figure 1. Schematic diagram of scAPAdb showing the pipeline of data pre-processing and poly(A) site identification, the back-end data of APA information stored in scAPAdb and functional modules in scAPAdb.

APA landscape organized by experiments or studies

We categorized the data in scAPAdb by experiments or by studies. An APA atlas of a study is compiled from multiple experiments of the same submission. In scAPAdb, the 'Experiment' page catalogues APA datasets by experiments, which can be filtered by different categories such as species, year of publish, sequencing protocol and tissue (Figure 2A). Datasets of different experiments or studies are displayed as individual dataset cards, and users can view detailed descriptive and statistical information of the chosen dataset or download APA data with full genome and cell type an-

notation. With scAPAtrap, 2617 to 31 335 3′ UTR poly(A) sites (median: 12 934 per experiment) were identified across experiments in these six species; with Sierra, 1398–23 284 3′ UTR poly(A) sites (median: 10,620 per experiment) were identified (Supplementary Table S1). The 'Study' page catalogues APA datasets by study, with 3491 to 47 068 3′ UTR poly(A) sites across studies (median: 16 577 per study) (Supplementary Table S2).

As an example, the mouse sperm scRNA-seq dataset was obtained from a previous study (12), which contains two experiments. Detailed APA information of this study, such

Table 1. Data summary in scAPAdb

Organism	Study#	Experiment#	Tissue#	Cell type#	Cell#
Homo sapiens	37	108	19	84	252 836
Mus musculus	36	192	49	129	511 366
Arabidopsis thaliana	9	50	2	27	220 736
Chlamydomonas reinhardtii	1	9	1	4	67 821
Oryza sativa Indica	1	1	1	8	12 326
Oryza sativa Japonica	1	1	1	8	10 965
Zea mays	1	2	1	2	13 507

as biological sample information, cell type annotation as well as profiles of gene expression, poly(A) site expression and APA usages in individual cells, is available by clicking the 'APA Details' button on the respective data card of the study (Figure 2B). This study contains 2042 cells in three cell types. Up to 19 558 3′ UTR poly(A) sites were compiled from the two experiments of this study. Global structure of the three cell types based on the gene-cell expression profile is shown by a UMAP plot (Figure 2C). The 2D embeddings of the two experiments were also displayed (Figure 2D).

APA usages and 3′ UTR length changes in single cells

scAPAdb provides two APA metrics, percentage of the proximal poly(A) site usage index (PPUI) and weighted 3′ UTR length (WUL), to quantify dynamic APA usages of a gene or all genes in individual cells. A smaller PPUI or higher WUL of a cell means potential global 3′ UTR lengthening (i.e. higher usage of the distal poly(A) site) in that cell. For instance, on the 'Experiment' page showing one sample of the mouse sperm study (12), the overall distribution of mean PPUI scores of all genes in individual cells was visualized by UMAP (Figure 2E) and a bar plot (Figure 2F), which reflects the gradual shortening of 3′ UTR in the transition from spermatocytes (SC) to elongating spermatids (ES).

Poly(A) signals of different cell types

We identified poly(A) signals for 3′ UTR poly(A) sites identified by scAPAtrap in individual cell types of each study. The DMEME tool (32) was adopted to search motifs significantly enriched within the near upstream (−50 to 0 nt), far upstream (−50 to −100 nt) and downstream (0 to 50 nt) regions of poly(A) sites. The identified motifs and single nucleotide compositions surrounding poly(A) sites of individual cell types are intuitively visualized (Figure 2G). Accordingly, detailed information of each motif is given, such as enriched matching words, the number of sequences matching the motif and *E*-value. Users can also download sequences surrounding poly(A) sites to a local computer for custom analysis.

Search a gene in scAPAdb

scAPAdb provides a very convenient search interface for querying gene(s) by commonly used keywords such as gene

identifier, gene symbol/alias, gene description, gene function, and gene ontology (GO) ID (Figure 2H). Users can also search an experiment or a study by an accession number. A query with gene relevant keywords will lead to an intermediate page showing the experiments with at least one poly(A) site identified in the queried gene(s) (Figure 2I). Then users can choose to view the full APA information of the queried gene in a selected experiment. A query with accession number will directly lead to the page of the quired experiment or study.

Take the *Odf4* gene for example, which has been reported to play an important role in sperm tails (33). After searching for this gene in the database, an intermediate page opens, providing a detailed gene list associated with the input keyword (Figure 2I). By clicking the gene of interest (here is ENSMUSG00000032921), a new page opens to show all experiments containing this gene. By clicking the dataset card of an experiment (here is GSM2803334), a new webpage appears to show detailed information of the gene and the usage of poly(A) sites of this gene across cell types. Two poly(A) sites were detected in the 3′ UTR of this gene by scAPAtrap, whereas only one poly(A) site was identified by Sierra (Figure 2J). The read coverage of the distal 3′ UTR poly(A) site in this gene is increased from ES to SC, suggesting APA dynamics of 3′ UTR shortening during mouse spermatogenesis from SC to ES (Figure 2K). APA dynamics across cell types quantified by the PPUI score also reflect the gradual 3′ UTR shortening in the transition from SC to ES (Figure 2L). The scatter plot shows the correlation between the profile of APA usage and gene expression (Figure 2M).

Batch download

For each study or experiment, different files are available for download to meet the users' needs. For each experiment, we provided several files for download, including the gene-cell expression matrix, metafiles recording the cell type annotation and poly(A) site annotation, the curated matrices of poly(A) sites, PPUI scores and WUL scores by using scAPAtrap and Sierra. For each study, we also provided the metafile and the APA matrices compiled from different experiments. In addition to the APA content, the full list of the reference genome assembly and genome annotation for each species, and relevant bioinformatics workflows used in scAPAdb are also described on the 'Download' page of scAPAdb.

SUMMARY AND FUTURE DIRECTIONS

We present a user-friendly database, scAPAdb, which provides comprehensive single-cell APA atlas in animals and plants. By collecting and processing massive 3′ tag-based scRNA-seq data from >360 experiments by two algorithms, scAPAdb compiles a compendium of poly(A) sites and APA events in individual cells of different cell types from different tissues and species. Rich information at single-cell resolution including the poly(A) site expression, APA usage, genome annotation, poly(A) signals, poly(A) sequences and profile of gene expression is present. scAPAdb also provides an easy-to-use web service for users to query the APA information of a single gene across the whole database. We will
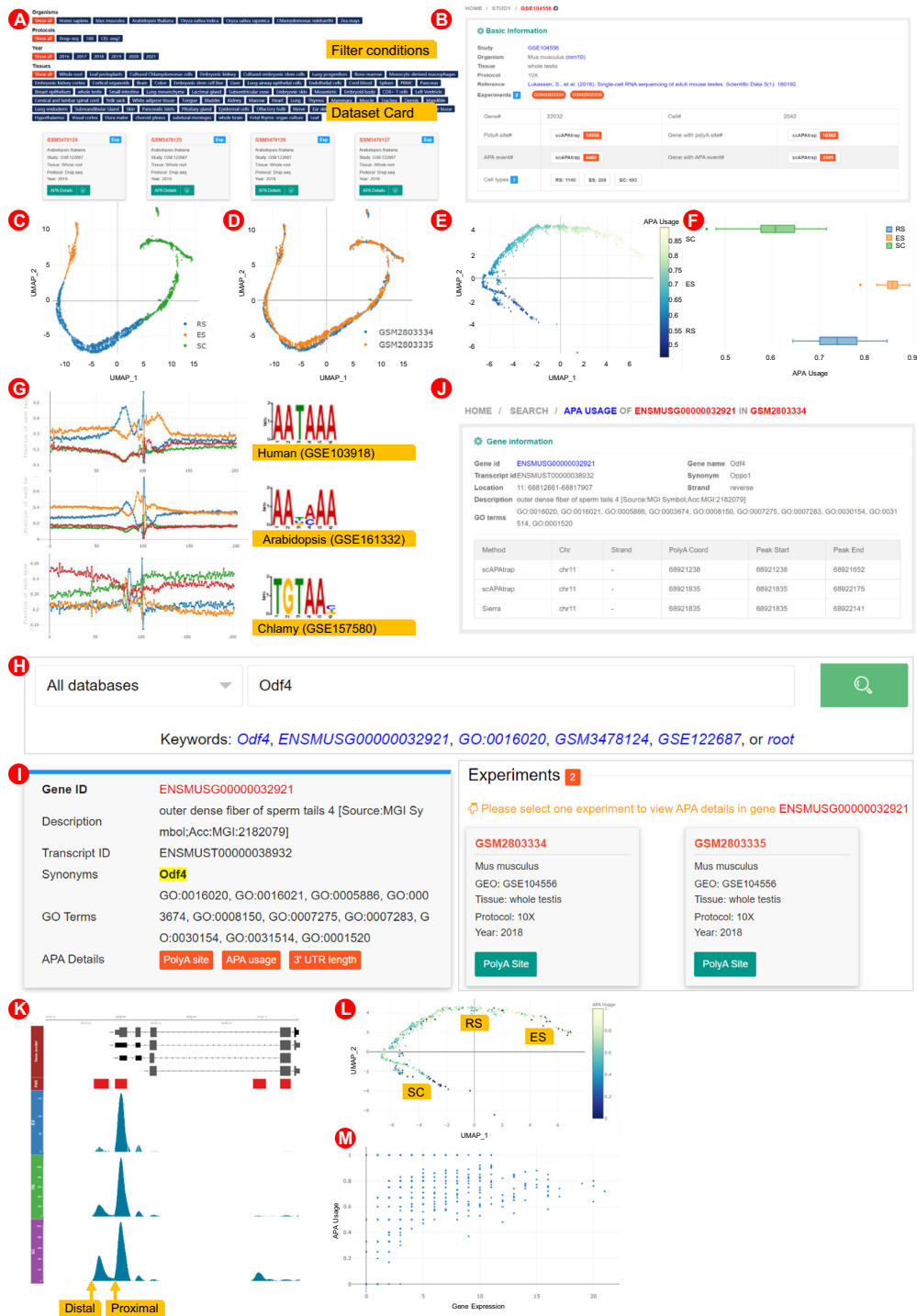
Figure 2. Exploration of the APA landscape with scAPAdb. The 'Experiment' page presents each experiment as a dataset card and users can filter experiments by entries including species, year of publish, sequencing protocol and tissue (**A**). Upon the selection of a study (here is GSE104556, access via http://www.bmibig.cn/scAPAdb/groups/Study/study_info.php?study=GSE104556), detailed APA information on this study is provided, including summary and statistics of this study (**B**), the global structure of cell type distribution shown by a UMAP plot (**C**) and 2D embeddings of all experiments in this study (**D**). Similarly, upon the selection of an experiment (access via http://www.bmibig.cn/scAPAdb/groups/Dataset/dataset_info.php?GSM=GSM2803334), the APA usage of individual cells can be visualized by an UMAP plot (**E**) or a bar plot (**F**). Single nucleotide compositions around poly(A) sites and poly(A) signal motifs of selected cell types are provided in the 'PolyA Signal' module (**G**). Users can easily search the entire database for a gene, an experiment or a study by a variety of keywords through the search interface (**H**). Upon the query of the keyword '*Odf4*', an intermediate page appears to show a gene list associated with the input keyword and users can click the gene of interest (here is ENSMUSG00000032921) to show all experiments with at least one poly(A) site expressed in this gene (**I**). By clicking the dataset card of the experiment GSM2803334 from the search results, detailed APA information of the gene is provided, including summary of the gene and APA sites (**J**), the read coverage of poly(A) sites along this gene across cell types (**K**), APA dynamics across cell types (**L**) and the scatter plot showing the correlation between profiles of APA usage measured by PPUI and gene expression (**M**). SC, spermatocytes; ES, elongating spermatids; RS, round spermatids; PPUI, percentage of the proximal poly(A) site usage index.

continue to update scAPAdb to add more species and more diverse cell/tissue types. At present, scAPAdb only focuses on 3′ tag-based scRNA-seq data; we expect to leverage the merit of full-length scRNA-seq (e.g. Smart-seq2) to expand our APA compendium when new efficient computational methods are available. The comprehensive APA atlas in single cells from various cell types, tissues and species provided in scAPAdb would help us to investigate cell-to-cell heterogeneity in APA isoform usages and gene expression, explore new marker genes, discover novel and/or rare cell types and elucidate APA dynamics and APA-mediated gene regulation at the single-cell level.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

REFERENCES

1. Tian,B. and Manley,J.L. (2017) Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell Biol.*, **18**, 18–30.
2. Gruber,A.J. and Zavolan,M. (2019) Alternative cleavage and polyadenylation in health and disease. *Nat. Rev. Genet.*, **20**, 1.
3. Tian,B. and Manley,J.L. (2013) Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem. Sci.*, **38**, 312–320.
4. Ji,G., Guan,J., Zeng,Y., Li,Q.Q. and Wu,X. (2015) Genome-wide identification and predictive modeling of polyadenylation sites in eukaryotes. *Brief. Bioinform.*, **16**, 304–313.
5. Chen,M., Ji,G., Fu,H., Lin,Q., Ye,C., Ye,W., Su,Y. and Wu,X. (2020) A survey on identification and quantification of alternative polyadenylation sites from RNA-seq data. *Brief. Bioinform.*, **21**, 1261–1276.
6. Ziegenhain,C., Vieth,B., Parekh,S., Reinius,B., Guillaumet-Adkins,A., Smets,M., Leonhardt,H., Heyn,H., Hellmann,I. and Enard,W. (2017) Comparative analysis of single-cell RNA sequencing methods. *Mol. Cell*, **65**, 631–643.
7. Ye,C., Zhou,Q., Hong,Y. and Li,Q.Q. (2019) Role of alternative polyadenylation dynamics in acute myeloid leukaemia at single-cell resolution. *RNA Biology*, **16**, 785–797.
8. Göpferich,M., George,N.O., Muelas,A.D., Bizyn,A., Pascual,R., Fijalkowska,D., Kalamakis,G., Müller,U., Krijgsveld,J., Mendez,R. *et al.* (2020) Single cell 3′UTR analysis identifies changes in alternative polyadenylation throughout neuronal differentiation and in autism. bioRxiv doi: https://doi.org/10.1101/2020.08.12.247627, 12 August 2020, preprint: not peer reviewed.
9. Cheng,L.C., Zheng,D., Baljinnyam,E., Sun,F., Ogami,K., Yeung,P.L., Hoque,M., Lu,C.-W., Manley,J.L. and Tian,B. (2020) Widespread transcript shortening through alternative polyadenylation in secretory cell differentiation. *Nat. Commun.*, **11**, 3182.
10. Agarwal,V., Lopez-Darwin,S., Kelley,D.R. and Shendure,J. (2021) The landscape of alternative polyadenylation in single cells of the developing mouse embryo. bioRxiv doi: https://doi.org/10.1101/2021.01.21.427498, 24 March 2021, preprint: not peer reviewed.
11. Levin,M., Zalts,H., Mostov,N., Hashimshony,T. and Yanai,I. (2020) Gene expression dynamics are a proxy for selective pressures on alternatively polyadenylated isoforms. *Nucleic Acids Res.*, **48**, 5926–5938.
12. Shulman,E.D. and Elkon,R. (2019) Cell-type-specific analysis of alternative polyadenylation using single-cell transcriptomics data. *Nucleic Acids Res.*, **47**, 10027–10039.
13. Patrick,R., Humphreys,D.T., Janbandhu,V., Oshlack,A., Ho,J.W.K., Harvey,R.P. and Lo,K.K. (2020) Sierra: discovery of differential transcript usage from polyA-captured single-cell RNA-seq data. *Genome Biol.*, **21**, 167.
14. Wu,X., Liu,T., Ye,C., Ye,W. and Ji,G. (2020) scAPAtrap: identification and quantification of alternative polyadenylation sites from single-cell RNA-seq data. *Brief. Bioinform.*, **22**, bbaa273.
15. Gao,Y., Li,L., Amos,C.I. and Li,W. (2021) Analysis of alternative polyadenylation from single-cell RNA-seq using scDaPars reveals cell subpopulations invisible to gene expression. *Genome Res.*, https://doi.org/10.1101/gr.271346.120.
16. Hashimshony,T., Wagner,F., Sher,N. and Yanai,I. (2012) CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep.*, **2**, 666–673.
17. Macosko,E.Z., Basu,A., Satija,R., Nemesh,J., Shekhar,K., Goldman,M., Tirosh,I., Bialas,A.R., Kamitaki,N., Martersteck,E.M. *et al.* (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, **161**, 1202–1214.
18. Zheng,G.X., Terry,J.M., Belgrader,P., Ryvkin,P., Bent,Z.W., Wilson,R., Ziraldo,S.B., Wheeler,T.D., McDermott,G.P., Zhu,J. *et al.* (2017) Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.*, **8**, 14049.
19. Müller,S., Rycak,L., Afonso-Grunz,F., Winter,P., Zawada,A.M., Damrath,E., Scheider,J., Schmäh,J., Koch,I., Kahl,G. *et al.* (2014) APADB: a database for alternative polyadenylation and microRNA regulation events. *Database*, **2014**, bau076.
20. You,L., Wu,J., Feng,Y., Fu,Y., Guo,Y., Long,L., Zhang,H., Luan,Y., Tian,P., Chen,L. *et al.* (2014) APASdb: a database describing alternative poly(A) sites and selection of heterogeneous cleavage sites downstream of poly(A) signals. *Nucleic. Acids. Res.*, **43**, D59–D67.
21. Wang,R., Nambiar,R., Zheng,D. and Tian,B. (2018) PolyA_DB 3 catalogs cleavage and polyadenylation sites identified by deep sequencing in multiple genomes. *Nucleic Acids Res.*, **46**, D315–D319.
22. Zhu,S., Ye,W., Ye,L., Fu,H., Ye,C., Xiao,X., Ji,Y., Lin,W., Ji,G. and Wu,X. (2020) PlantAPAdb: a comprehensive database for alternative polyadenylation sites in plants. *Plant Physiol.*, **182**, 228–242.
23. Herrmann,C.J., Schmidt,R., Kanitz,A., Artimo,P., Gruber,A.J. and Zavolan,M. (2020) PolyASite 2.0: a consolidated atlas of polyadenylation sites from 3′ end sequencing. *Nucleic Acids Res.*, **48**, D174–D179.
24. Feng,X., Li,L., Wagner,E.J. and Li,W. (2018) TC3A: the cancer 3′ UTR atlas. *Nucleic Acids Res.*, **46**, D1027–D1030.
25. Jin,W., Zhu,Q., Yang,Y., Yang,W., Wang,D., Yang,J., Niu,X., Yu,D. and Gong,J. (2021) Animal-APAdb: a comprehensive animal alternative polyadenylation database. *Nucleic. Acids. Res.*, **49**, D47–D54.
26. Marini,F., Scherzinger,D. and Danckwardt,S. (2020) TREND-DB—a transcriptome-wide atlas of the dynamic landscape of alternative polyadenylation. *Nucleic. Acids. Res.*, **49**, D243–D253.
27. Hong,W., Ruan,H., Zhang,Z., Ye,Y., Liu,Y., Li,S., Jing,Y., Zhang,H., Diao,L., Liang,H. *et al.* (2020) APAatlas: decoding alternative polyadenylation across human tissues. *Nucleic. Acids. Res.*, **48**, D35–D39.
28. Ha,K.C.H., Blencowe,B.J. and Morris,Q. (2018) QAPA: a new method for the systematic analysis of alternative polyadenylation from RNA-seq data. *Genome Biol.*, **19**, 45.
29. Xia,Z., Donehower,L.A., Cooper,T.A., Neilson,J.R., Wheeler,D.A., Wagner,E.J. and Li,W. (2014) Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3′- UTR landscape across seven tumour types. *Nat. Commun.*, **5**, 5274–5274.
30. Franzén,O., Gan,L.-M. and Björkegren,J.L.M. (2019) PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database*, **2019**, baz046.
31. Ye,W., Liu,T., Fu,H., Ye,C., Ji,G. and Wu,X. (2020) movAPA: modeling and visualization of dynamics of alternative polyadenylation across biological samples. *Bioinformatics*, **37**, 2470–2472.
32. Bailey,T.L. (2011) DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics*, **27**, 1653–1659.
33. Ghafouri-Fard,S., Ghafouri-Fard,S. and Modarressi,M.H. (2012) Expression of splice variants of cancer-testis genes ODF3 and ODF4 in the testis of a prostate cancer patient. *Genet. Mol. Res.: GMR*, **11**, 3642–3648.