



HAL
open science

MatrixDB 2024: an increased coverage of extracellular matrix interactions, a new Network Explorer and a new web interface

Kasun W Samarasinghe, Max Kotlyar, Sylvain D Vallet, Catherine Hayes, Alexandra Naba, Igor Jurisica, Frédérique Lisacek, Sylvie Ricard-Blum

► To cite this version:

Kasun W Samarasinghe, Max Kotlyar, Sylvain D Vallet, Catherine Hayes, Alexandra Naba, et al.. MatrixDB 2024: an increased coverage of extracellular matrix interactions, a new Network Explorer and a new web interface. *Nucleic Acids Research*, 2025, 53 (D1), pp.D1677 - D1682. 10.1093/nar/gkae1088 . hal-04869478

HAL Id: hal-04869478

<https://hal.science/hal-04869478v1>

Submitted on 7 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

MatrixDB 2024: an increased coverage of extracellular matrix interactions, a new Network Explorer and a new web interface

Kasun W. Samarasinghe¹, Max Kotlyar², Sylvain D. Vallet³, Catherine Hayes¹,
Alexandra Naba⁴, Igor Jurisica^{2,5,6}, Frédérique Lisacek^{1,*} and Sylvie Ricard-Blum^{7,*}

¹SIB, Swiss Institute of Bioinformatics, Geneva, Switzerland

²Osteoarthritis Research Program, Division of Orthopedic Surgery, Schroeder Arthritis Institute and Data Science Discovery Centre for Chronic Diseases, Krembil Research Institute, University Health Network, Toronto, ON M5T 0S8, Canada

³Institut de Biologie Structurale, UMR 5075, CEA, CNRS, Université Grenoble Alpes, Grenoble 38000, France

⁴Department of Physiology and Biophysics, University of Illinois Chicago, Chicago, IL 60612, USA

⁵Departments of Medical Biophysics and Computer Science, and the Faculty of Dentistry, University of Toronto, Toronto, Ontario, Canada

⁶Institute of Neuroimmunology, Slovak Academy of Sciences, Bratislava, Slovakia

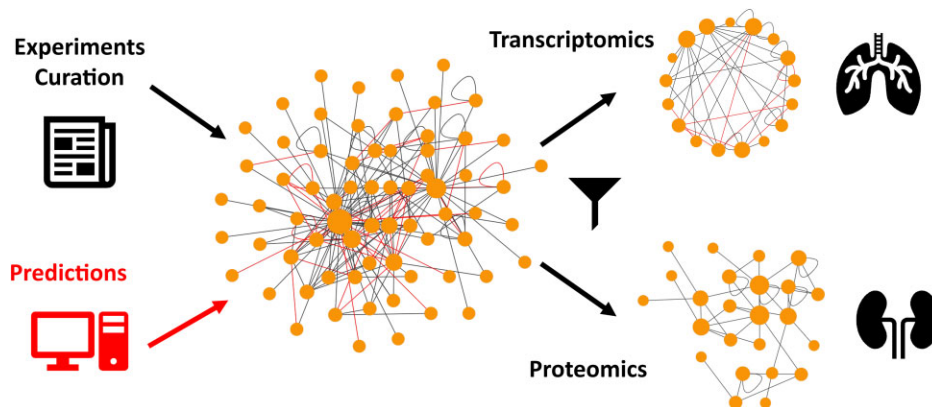
⁷Institut de Chimie et Biochimie Moléculaires et Supramoléculaires (ICBMS), UMR 5246, CNRS, Université Lyon 1, Villeurbanne 69622, France

*To whom correspondence should be addressed. Tel: +33 478244232; Email: sylvie.ricard-blum@univ-lyon1.fr
Correspondence may also be addressed to Frédérique Lisacek. Email: frederique.lisacek@sib.swiss

Abstract

MatrixDB, a member of the International Molecular Exchange consortium (IMEx), is a curated interaction database focused on interactions established by extracellular matrix (ECM) constituents including proteins, proteoglycans, glycosaminoglycans and ECM bioactive fragments. The architecture of MatrixDB was upgraded to ease interaction data export, allow versioning and programmatic access and ensure sustainability. The new version of the database includes more than twice the number of manually curated and experimentally-supported interactions. High-confidence predicted interactions were imported from the Integrated Interactions Database to increase the coverage of the ECM interactome. ECM and ECM-associated proteins of five species (human, murine, bovine, avian and zebrafish) were annotated with matrisome divisions and categories, which are used for computational analyses of ECM -omic datasets. Biological pathways from the Reactome Pathway Knowledgebase were also added to the biomolecule description. New transcriptomic and expanded proteomic datasets were imported in MatrixDB to generate cell- and tissue-specific ECM networks using the newly developed in-house Network Explorer integrated in the database. MatrixDB is freely available at <https://matrixdb.univ-lyon1.fr>.

Graphical abstract



Introduction

MatrixDB is a curated database focused on interactions mediated by the components of the extracellular matrix (ECM), namely proteins, proteoglycans, glycosaminoglycans, and bioactive ECM fragments collectively referred to as ma-

trikines or matricryptins (1). One of the specific features of MatrixDB is the curation of ECM interactions at several molecular levels, namely matrikines/matricryptins (bioactive fragments, e.g. endostatin), individual polypeptide chains (e.g. collagen α chains), and native multimeric ECM molecules

Received: September 15, 2024. Revised: October 23, 2024. Editorial Decision: October 24, 2024. Accepted: October 27, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

(e.g. trimeric collagens). The database is freely available through a web interface (<https://matrixdb.univ-lyon1.fr>), does not require any login or registration and is not password-protected. MatrixDB has been regularly updated since its creation in 2009 (2–5). MatrixDB is one of the services of the French node of ELIXIR (<https://elixir-europe.org/about-us/who-we-are/nodes/france>), and an active member of the International Molecular Exchange (IMEx) consortium (<https://www.imexconsortium.org/>) (6), which is a Global Core Biodata Resource (<https://globalbiodata.org>). Interaction data available in MatrixDB are manually curated from the literature via the curation interface of the IntAct database (7) following the IMEx curation rules (8).

The architecture of MatrixDB was designed to improve reusability (ease of data export and implementation of programmatic access) and long-term sustainability (allowing for versioning, and strongly limiting dependencies to unsupported technology), and to automate data import from an extended number of sources as detailed below. The content of the database was expanded, and MatrixDB contains twice as many experimentally-supported and manually curated interactions as the previous release (5). High-confidence predicted interactions from the Integrated Interactions Database (IID) (9), were added to increase the coverage of the ECM interaction networks generated with the new in-house ‘Network Explorer’ tool integrated in the database. Matrisome divisions and categories (10), now a standard in computational analyses of ECM -omic datasets (11,12), were added to annotate ECM and ECM-associated proteins of five species together with biological pathways imported from the Reactome Pathway Knowledgebase (13). Transcriptomic data and an expanded ECM proteomic dataset were imported from Bgee, an integrated curated expression atlas (14), and MatrisomeDB 2.0 (15) respectively to increase the coverage of cell- and tissue-specific interaction networks generated with the Network Explorer. Selected interaction data can now be exported in a tabular format and the corresponding networks in Cytoscape format in addition to image capture for further analyses.

Architecture and management system of the database

The new system architecture comprises two main subsystems, data integration and the Web portal (see [Supplementary Figure S1, Supplementary Methods, https://github.com/glyco-expasy/matrixdb-api](#) and <https://github.com/glyco-expasy/matrixdb-ui> for details). This version (release 4.0) is and the future ones will be archived in the same GitHub repository. The database content was shifted from the unsupported ACEDB database manager software (16) to a NoSQL design (MongoDB, <https://www.mongodb.com>). A simple data model was defined to represent the entities of MatrixDB, which ideally could extend to other formats/models such as RDF (Resource Description Framework). It is implemented using MongoDB v.6.0, a NoSQL database.

Data integration and visualization

The database was built by integrating data from several open biological data sources as detailed below. This integration process followed the Extract Transform Load (ETL) pattern, where data were extracted from external Application Programming Interfaces (APIs) or other forms of distributions

(zip archives), then transformed according to the data model and loaded to the database. The data integration framework was developed in Python to perform the ETL process, considering the increased volume of data to be processed and the variety of input data formats. For each external data source, a data integration pipeline consisting of independent ETL sub-components was developed, with parallel processing, whenever necessary to handle large data volumes. All pipelines were designed to maximize reusability in the data integration process, and ease regular updates of MatrixDB content. Third-party JavaScript libraries, Cytoscape (17), Mol* (18) and D3 (<https://d3js.org/>) were used to implement visual components ([Supplementary Table S1](#)).

An expanded interaction dataset

Experimentally-supported interaction data

The new interaction dataset (release 4.0) comprises experimentally-supported interactions manually curated from publications by MatrixDB, and interactions involving at least one ECM component imported from IntAct (release 2024-02) (7), and other databases of the IMEx consortium (6). Orthologous experimental protein-protein interactions are not inferred to human species in the new version of MatrixDB. The new experimental interaction dataset comprises ~243 000 interactions, a two-fold increase in the number of manually curated and experimentally-supported interaction data compared to the previous version (5) (available at <http://v1.matrixdb.univ-lyon1.fr/>) (Figure 1). This increase is mainly due to the import from IntAct database of large protein-protein interaction datasets including tens of thousands of interactions, some of them involving ECM proteins, such as BioPlex 2.0 (19), resulting in a 2.5-fold increase in the number of protein-protein interaction data. However, there is a ~4-fold increase in the number of GAG-protein interactions, from 294 to 1230, in release 4.0 (Figure 1) with a 3-fold increase in the number of heparin interactions (from 251 to 766). There is also a significant increase in the number of interactions of ECM proteins (e.g. a 4.4-fold increase in the number of interactions of lysyl oxidase-like 2, from 12 to 53 in the 4.0 release). MatrixDB curated dataset and the entire interaction dataset available in MatrixDB can be downloaded from the ‘Download’ tab under the MITAB 2.7 format.

Predicted protein–protein interactions

High-confidence computationally predicted protein–protein interactions (151 132) from the Integrated Interaction Database (9), a member of the IMEx consortium, were added to MatrixDB. These high-confidence interactions (see [Supplementary Methods](#) for details) were predicted by state-of-the-art computational methods (20–23). They are clearly labeled as such in the database, and can be included, or not, in the networks generated with the Network Explorer.

Display of interaction data

Interaction data are displayed in the ‘Biomolecule’ page as an interactive table in a first tab (Interaction list view), and an interactive network in a second tab (Network view). The number of experimentally supported and predicted interactions of the biomolecule is indicated at the top. The table comprises six columns including the names and identifiers of the participants, the nature of the interaction (direct or indirect), the

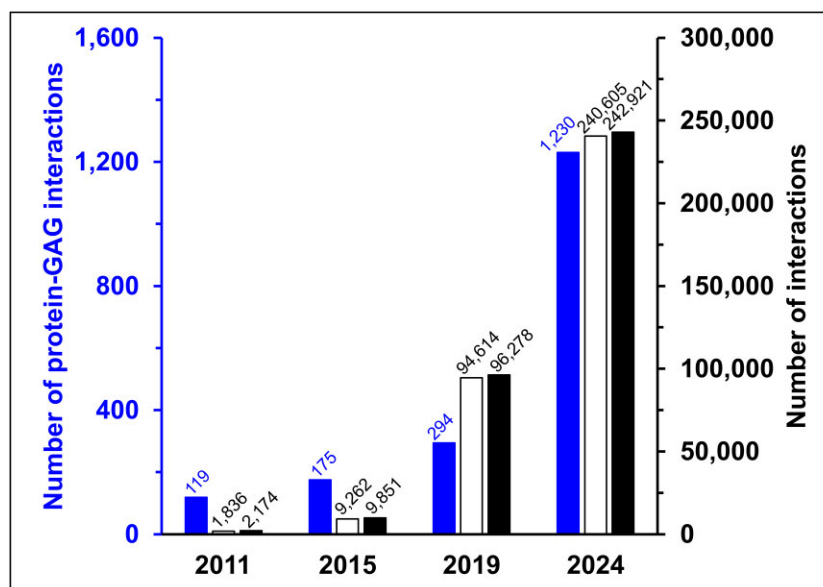


Figure 1. Number of experimentally-supported interactions in MatrixDB releases. The plot shows the increase in the total number of experimental interactions (black), and in the number of experimental protein-protein (white) and protein-GAG (blue) interactions. The ‘protein’ category includes monomeric proteins and individual polypeptide chains of multimers (e.g. α chains of collagens).

molecular interaction (MI) score imported from the IntAct database (7) for experimental interactions, and the type of interaction (experimental interactions in black or predicted interactions in red). Clicking on one interaction opens an ‘Association’ page that provides the identifiers of the participants, the MI score and the list of experiments supporting the interaction. Clicking on one experiment opens an ‘Experiment’ page where the interaction detection method, the participant detection method, the binding region, the stoichiometry, kinetics (association and dissociation rates), the equilibrium constant (affinity), and the effect of mutations on the interaction are displayed when available.

Annotation of biomolecules

Biological pathways were imported from the Reactome Pathway Knowledgebase (13), and displayed at the bottom of the ‘Biomolecule’ page together with Gene Ontology (GO) terms (24,25) and UniProtKB keywords (26). ECM and ECM-associated proteins were first annotated with matrisome divisions (core matrisome and matrisome-associated), and then with one of the six matrisome categories (ECM glycoproteins, collagens and proteoglycans for core matrisome components, and ECM-affiliated, ECM regulators, and secreted factors for matrisome-associated components), which are now broadly adopted for -omic analyses of ECM datasets. The following lists were used for matrisome annotations: human and murine matrisomes (27), bovine matrisome (28), avian matrisome (29) and zebrafish matrisome (30). Matrisome annotations are displayed on the upper right corner of the ‘Biomolecule’ page. An additional, manual, annotation called ‘MatrixDB ECM’ includes proteins, which are not annotated as matrisome proteins, but are annotated with at least one of selected UniProtKB keywords (Extracellular space/secreted: KW-0964, Extracellular matrix: KW-0272, and Basement membrane: KW-0084) or GO term Cellular Component (Extracellular space: GO:0005615, Extracellular region: GO:0005576, Extracellular matrix: GO:0031012,

and Basement membrane: GO:0005604). The list of ‘MatrixDB ECM’ proteins is available in the ‘Download’ tab.

Basic search and advanced search

Indexing and searches are performed with a configured version of Solr v.9.5 (<https://solr.apache.org/>), an open-source community software. Basic search queries all fields of the database, and includes free text, biomolecule and gene names, accession numbers of UniProtKB (26), Complex Portal (31) and Chemical Entities of Biological Interest (ChEBI) (32) and PMID identifiers. By default, the ten most relevant results (award icon, Supplementary Figure S2) are displayed where relevance is first defined by the field of the database where the searched term is found. For example, the term found in a biomolecule name is considered more relevant than in an annotation (see Supplementary Methods for further details). The eye icon allows users to display all the results (ten items per page), whereas the filter icon allows them to select a biomolecule category and/or a species (Supplementary Figure S2). Advanced search allows users to retrieve biomolecules using their names, identifiers, UniProtKB keywords, GO terms, matrisome divisions and categories, and Reactome biological pathways. The database can also be queried using a combination of species AND GO term or species AND matrisome division/category. The Biomolecule search displays biomolecules included in the database whether they have binding partners or not.

Transcriptomic and proteomic data

Transcriptomic data were imported from the Bgee expression atlas (14), comprising curated and re-annotated GTEx (Genotype Tissue Expression) samples based on healthy expression data (14) to replace GTEx data formerly integrated in MatrixDB. An expanded and curated ECM proteomic dataset including over 150 000 protein entries from 42 studies on the ECM on over 25 different healthy and diseased tissues and

organs was imported from MatrisomeDB 2.0 (15). These experimental datasets increase the coverage of cells and tissues used to filter the networks generated with the Network Explorer (see below). Transcriptomic data (transcripts per million) are displayed in anatomograms imported from Gene Expression Atlas (<https://www.ebi.ac.uk/gxa>) in a first tab and proteomic data (normalized spectral abundance factor, NSAF) as heatmaps in a second tab.

Network Explorer

User-selected biomolecules can be imported in the Network Explorer interface from the 'Biomolecule' page by clicking on the '+' button next to the green Network Explorer icon in the header. Once the biomolecule is added to the Explorer, the icon turns from green to white and clicking on it gives access to the Explorer. Biomolecules can also be imported from the 'Network Explorer' page, and the top window can be used to search the database. In contrast to the previous navigator of MatrixDB, the Network Explorer includes by default interactions between partners. They can be excluded by ticking the 'Exclude partner interactions' box. The Network Explorer remembers the selection of biomolecules in the browser session. Filters are available as a right menu to select network participants (e.g. those annotated as ECM biomolecule, in either MatrixDB ECM or matrisome lists), and interactions based on experimental and/or predicted interactions, expression level (transcriptomic data) or protein abundance (proteomic data). An additional filter allows user to select the interaction detection method. Filters can be reversed to display the unfiltered network. In addition to an option for image capture (camera icon), selected interaction data can be exported in tabular format (table icon) and the corresponding networks in Cytoscape format (Cytoscape icon) for further analyses. The search in the Network Explorer is restricted to biomolecules which are involved in at least one interaction.

Interaction networks including various proteoforms of ECM proteins

One of the specific features of MatrixDB is to ease the retrieval of interactions curated at several molecular levels and to visualize them as interaction networks reflecting what happens *in vivo*. This is illustrated in Figure 2, which recapitulates interactions established by several native trimeric collagen molecules (purple), which are rarely included in interactomes, individual collagen α chains (orange), collagen bioactive fragments (matrikines/matricryptins, yellow) and glycosaminoglycans (blue).

Increased coverage of ECM protein networks and tissue-specific networks

The interaction network of human fibronectin generated with the Network Explorer highlights its increased coverage due to the addition of predicted interactions. There is a 2-fold increase in the size of the fibronectin network upon addition of predicted interactions (638) to experimental ones (512) (Figure 3A and B). Filtering the entire fibronectin network using proteomic data at the tissue and NSAF levels results in fibronectin networks of various size depending on the tissue (e.g. blood vessel and eye, Figure 3C and D, respectively).

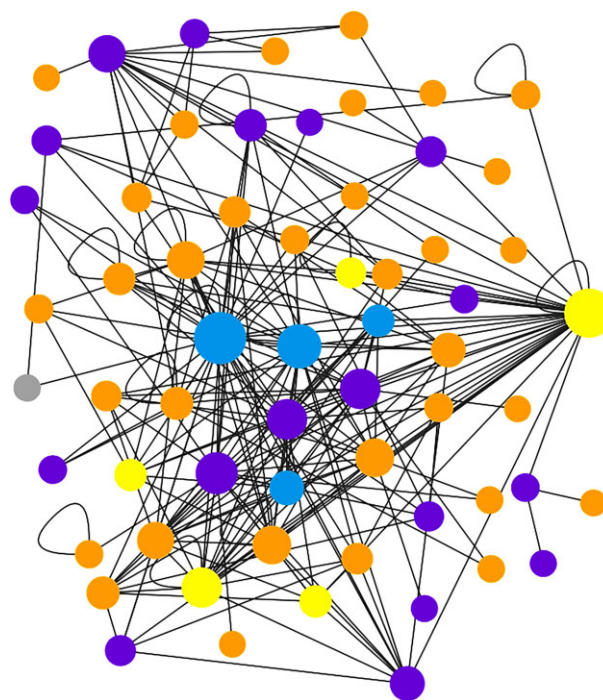


Figure 2. Interaction networks of several molecular proteoforms of human collagens. Network of selected native trimeric collagens, individual collagen α chains and collagen matrikines/matricryptins. It has been built using the Network Explorer using 'Collagen', 'Multimers' and '*Homo sapiens*' as filters. The biomolecules are color-coded (protein: orange, glycosaminoglycan (GAG): blue, multimer: purple, protein fragment (PFRAG): yellow, grey: small molecule).

Conclusion

The architecture and the management system of MatrixDB database were entirely rebuilt to replace the previous version, which was no longer supported, and ease data import and export as well as updates of MatrixDB content. Furthermore, the system architecture is designed with scalability as a core consideration, addressing both data integration and the web portal application in which a high volume of concurrent requests can be handled via the API framework, thereby accommodating an expanding database. The new version of MatrixDB includes a 2-fold expanded experimental dataset as well as predicted protein-protein interactions, curated transcriptomic data, an expanded ECM proteomic dataset, a new Network Explorer, and several export modes of user-selected interactions for further computational analysis.

On-going work aims to complete the curation of the glycosaminoglycan interaction dataset, MatrixDB being the primary database to specifically curate and report those interactions. We will also pursue the curation of low-throughput, biophysical, interactions of native ECM molecules and matrikines/matricryptins to build a comprehensive ECM interaction network including them. Mutations affecting interactions and associated with rare ECM genetic disorders (e.g. collagenopathies such as Ehlers-Danlos syndrome, osteogenesis imperfecta or epidermolysis bullosa) (33), will be mapped on tissue- and cell-specific ECM networks to determine how these networks are rewired by these mutations, and to decipher the molecular mechanisms altering ECM organization and function in these diseases.

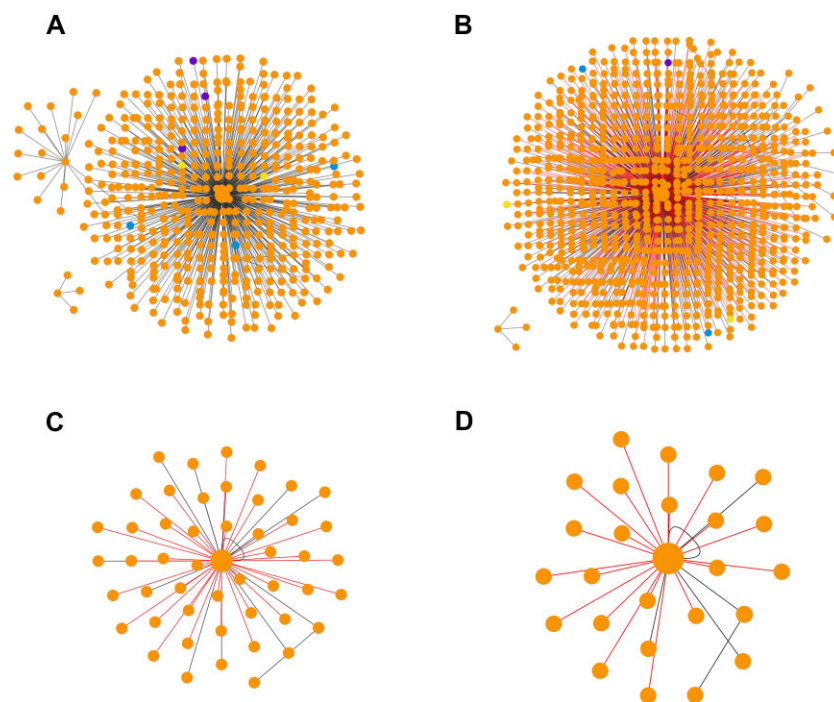


Figure 3. Interaction networks of human fibronectin generated using the Network Explorer of MatrixDB. **(A)** Fibronectin network comprising experimental interactions only. The small network located in the bottom left corner comprises an isoform of fibronectin (P02751–12) and its three partners. The biomolecules are color-coded (protein: orange, glycosaminoglycan (GAG): blue, multimer: purple, protein fragment (PFRAG): yellow). **(B)** Fibronectin network comprising both experimental (black links) and predicted interactions (red links). **(C)** Fibronectin network in blood vessel generated by filtering the entire network with proteomic data (normalized spectral abundance factor 501). **(D)** Fibronectin network in the eye generated by filtering the entire network with proteomic data (normalized spectral abundance factor 501).

Data availability

MatrixDB is freely available at <https://matrixdb.univ-lyon1.fr>. It is distributed under the terms of the Creative Commons Attribution Licence CC BY 4.0.

Supplementary Data

[Supplementary Data](#) are available at NAR Online.

Acknowledgements

We gratefully acknowledge the key contribution of Nicolas Thierry-Mieg (TIMC, University Grenoble Alpes, France) to the creation and previous updates of MatrixDB, and for his very valuable comments and suggestions to improve the new release. We also warmly thank Asia Paguntalan (University of Illinois, Chicago, USA) for beta-testing the new MatrixDB release and her suggestions, and Coline Berthollier (University Lyon 1) for her contribution to the curation of interactions.

Author contributions: K.S.: Methodology, design and implementation of MatrixDB new portal website and pipelines for data, import of data from several sources, review and editing. M.K.: Prediction of protein-protein interactions, review and editing. S.D.V.: Curation of interaction data, figure design and drawing, review & editing. C.H.: contribution to data model and quality control of pipelines, review & editing. A.N.: Provider of proteomic datasets, writing, review & editing. I.J.: Prediction of protein-protein interactions, writing, review & editing. F.L.: Conceptualization, methodology, writing, review & editing, funding. S.R.-B.: Conceptualization, curation of interaction data, contribution to the design of

MatrixDB new portal website, data import and display, drafting the manuscript, writing, figure design and drawing, review & editing, funding.

Funding

Swiss National Science Foundation, Switzerland [SNSF_310 030_208 178 to K.S., F.L., S.R.-B.]; Institut National du Cancer, France [INCa N° 2020-123 TEN-MAX to S.R.-B.]; the analysis and processing of ECM proteomic datasets integrated in MatrixDB was supported by the National Human Genome Research Institute (NHGRI) of the National Institutes of Health and the National Institutes of Health Common Fund through the Office of Strategic Coordination/Office of the NIH Director [U01HG012680 to A.N.]; National Cancer Institute [R21CA261642 to A.N.]. Funding for open access charge: [SNSF_310 030_208 178 to K.S., F.L., S.R.-B.]; INCa [2020-123 TEN-MAX to S.R.-B.].

Conflict of interest statement

None declared.

References

- Ricard-Blum,S. and Vallet,S.D. (2019) Fragments generated upon extracellular matrix remodeling: biological regulators and potential drugs. *Matrix Biol.*, 75–76, 170–189.
- Chautard,E., Ballut,L., Thierry-Mieg,N. and Ricard-Blum,S. (2009) MatrixDB, a database focused on extracellular protein-protein and protein-carbohydrate interactions. *Bioinformatics*, 25, 690–691.

3. Chautard,E., Fatoux-Ardore,M., Ballut,L., Thierry-Mieg,N. and Ricard-Blum,S. (2011) MatrixDB, the extracellular matrix interaction database. *Nucleic Acids Res.*, **39**, D235–D240.
4. Launay,G., Salza,R., Multedo,D., Thierry-Mieg,N. and Ricard-Blum,S. (2015) MatrixDB, the extracellular matrix interaction database: updated content, a new navigator and expanded functionalities. *Nucleic Acids Res.*, **43**, D321–D327.
5. Clerc,O., Deniaud,M., Vallet,S.D., Naba,A., Rivet,A., Perez,S., Thierry-Mieg,N. and Ricard-Blum,S. (2019) MatrixDB: integration of new data with a focus on glycosaminoglycan interactions. *Nucleic Acids Res.*, **47**, D376–D381.
6. Porras,P., Barrera,E., Bridge,A., Del-Toro,N., Cesareni,G., Duesbury,M., Hermjakob,H., Iannuccelli,M., Jurisica,I., Kotlyar,M., *et al.* (2020) Towards a unified open access dataset of molecular interactions. *Nat. Commun.*, **11**, 6144.
7. Del Toro,N., Shrivastava,A., Ragueneau,E., Meldal,B., Combe,C., Barrera,E., Perfetto,L., How,K., Ratan,P., Shirodkar,G., *et al.* (2022) The IntAct database: efficient access to fine-grained molecular interaction data. *Nucleic Acids Res.*, **50**, D648–D653.
8. Orchard,S., Kerrien,S., Abbani,S., Aranda,B., Bhate,J., Bidwell,S., Bridge,A., Briganti,L., Brinkman,F.S.L., Brinkman,F., *et al.* (2012) Protein interaction data curation: the International Molecular Exchange (IMEx) consortium. *Nat. Methods*, **9**, 345–350.
9. Kotlyar,M., Pastrello,C., Ahmed,Z., Chee,J., Varyova,Z. and Jurisica,J. (2022) IID 2021: towards context-specific protein interaction analyses by increased coverage, enhanced annotation and enrichment analysis. *Nucleic Acids Res.*, **50**, D640–D647.
10. Hynes,R.O. and Naba,A. (2012) Overview of the matrisome—an inventory of extracellular matrix constituents and functions. *Cold Spring Harb. Perspect. Biol.*, **4**, a004903.
11. Naba,A., Clauser,K.R., Ding,H., Whittaker,C.A., Carr,S.A. and Hynes,R.O. (2016) The extracellular matrix: tools and insights for the ‘omics’ era. *Matrix Biol.*, **49**, 10–24.
12. Naba,A. and Ricard Blum,S. (2020) The extracellular matrix goes -omics: resources and tools. In: *Extracellular Matrix Omics*. Biology of Extracellular Matrix. Springer, Vol. 7, pp. 1–16.
13. Milacic,M., Beavers,D., Conley,P., Gong,C., Gillespie,M., Griss,J., Haw,R., Jassal,B., Matthews,L., May,B., *et al.* (2024) The Reactome Pathway Knowledgebase 2024. *Nucleic Acids Res.*, **52**, D672–D678.
14. Bastian,F.B., Roux,J., Niknejad,A., Comte,A., Fonseca Costa,S.S., de Farias,T.M., Moretti,S., Parmentier,G., de Laval,V.R., Rosikiewicz,M., *et al.* (2021) The Bgee suite: integrated curated expression atlas and comparative transcriptomics in animals. *Nucleic Acids Res.*, **49**, D831–D847.
15. Shao,X., Gomez,C.D., Kapoor,N., Considine,J.M., Grams,C., Gao,Y.T. and Naba,A. (2023) MatrisomeDB 2.0: 2023 updates to the ECM-protein knowledge database. *Nucleic Acids Res.*, **51**, D1519–D1530.
16. Durbin,R. and Thierry-Mieg,J. (1994) The ACEDB genome database. In: *Computational Methods in Genome Research*. Springer Science & Business Media, NY, pp. 45–55.
17. Franz,M., Lopes,C.T., Huck,G., Dong,Y., Sumer,O. and Bader,G.D. (2016) Cytoscape.js: a graph theory library for visualisation and analysis. *Bioinformatics*, **32**, 309–311.
18. Sehnal,D., Bittrich,S., Deshpande,M., Svobodová,R., Berka,K., Bazziger,V., Velankar,S., Burley,S.K., Koča,J. and Rose,A.S. (2021) Mol* viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Res.*, **49**, W431–W437.
19. Huttlin,E.L., Bruckner,R.J., Paulo,J.A., Cannon,J.R., Ting,L., Baltier,K., Colby,G., Gebreab,F., Gygi,M.P., Parzen,H., *et al.* (2017) Architecture of the human interactome defines protein communities and disease networks. *Nature*, **545**, 505–509.
20. Kotlyar,M., Pastrello,C., Pivetta,F., Lo Sardo,A., Cumbaa,C., Li,H., Naranian,T., Niu,Y., Ding,Z., Vafaee,F., *et al.* (2015) In silico prediction of physical protein interactions and characterization of interactome orphans. *Nat. Methods*, **12**, 79–84.
21. Rhodes,D.R., Tomlins,S.A., Varambally,S., Mahavisno,V., Barrette,T., Kalyana-Sundaram,S., Ghosh,D., Pandey,A. and Chinnaiyan,A.M. (2005) Probabilistic model of the human protein-protein interaction network. *Nat. Biotechnol.*, **23**, 951–959.
22. Elefsinioti,A., Saraç,Ö.S., Hegele,A., Plake,C., Hubner,N.C., Poser,I., Sarov,M., Hyman,A., Mann,M., Schroeder,M., *et al.* (2011) Large-scale de novo prediction of physical protein-protein association. *Mol. Cell. Proteomics*, **10**, M111.010629.
23. Zhang,Q.C., Petrey,D., Deng,L., Qiang,L., Shi,Y., Thu,C.A., Bisikirska,B., Lefebvre,C., Accili,D., Hunter,T., *et al.* (2012) Structure-based prediction of protein-protein interactions on a genome-wide scale. *Nature*, **490**, 556–560.
24. Ashburner,M., Ball,C.A., Blake,J.A., Botstein,D., Butler,H., Cherry,J.M., Davis,A.P., Dolinski,K., Dwight,S.S., Eppig,J.T., *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.
25. Gene Ontology Consortium, Aleksander,S.A., Balhoff,J., Carbon,S., Cherry,J.M., Drabkin,H.J., Ebert,D., Feuermann,M., Gaudet,P., Harris,N.L., *et al.* (2023) The gene ontology knowledgebase in 2023. *Genetics*, **224**, iyad031.
26. UniProt Consortium (2023) UniProt: the Universal Protein knowledgebase in 2023. *Nucleic Acids Res.*, **51**, D523–D531.
27. Naba,A., Clauser,K.R., Hoersch,S., Liu,H., Carr,S.A. and Hynes,R.O. (2012) The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Mol. Cell. Proteomics*, **11**, M111.014647.
28. Listrat,A., Boby,C., Tournayre,J. and Jousse,C. (2023) Bovine extracellular matrix proteins and potential role in meat quality: first *insilico*Bostauruscompendium. *J. Proteomics*, **279**, 104891.
29. Huss,D.J., Saias,S., Hamamah,S., Singh,J.M., Wang,J., Dave,M., Kim,J., Eberwine,J. and Lansford,R. (2019) Avian primordial germ cells contribute to and interact with the extracellular matrix during early migration. *Front. Cell Dev. Biol.*, **7**, 35.
30. Nauroy,P., Hughes,S., Naba,A. and Ruggiero,F. (2018) The in-silico zebrafish matrisome: a new tool to study extracellular matrix gene and protein functions. *Matrix Biol.*, **65**, 5–13.
31. Meldal,B.H.M., Perfetto,L., Combe,C., Lubiana,T., Ferreira Cavalcante,J.V., Bye-A-Jee,H., Waagmeester,A., Del-Toro,N., Shrivastava,A., Barrera,E., *et al.* (2022) Complex Portal 2022: new curation frontiers. *Nucleic Acids Res.*, **50**, D578–D586.
32. Hastings,J., Owen,G., Dekker,A., Ennis,M., Kale,N., Muthukrishnan,V., Turner,S., Swainston,N., Mendes,P. and Steinbeck,C. (2016) ChEBI in 2016: improved services and an expanding collection of metabolites. *Nucleic Acids Res.*, **44**, D1214–D1219.
33. Lamandé,S.R. and Bateman,J.F. (2020) Genetic disorders of the extracellular matrix. *Anat Rec (Hoboken)*, **303**, 1527–1542.