OGP: A Repository of Experimentally Characterized O-Glycoproteins to Facilitate Studies on O-Glycosylation

Jiang-Ming Huang[1,2,#,a], Meng-Xi Wu[1,2,#,b], Yang Zhang[1,2,c], Si-Yuan Kong[1,d], Ming-Qi Liu[1,e], Bi-Yun Jiang[1,2,f], Peng-Yuan Yang[1,2,3,*,g], Wei-Qian Cao[1,2,3,*,h]

[1] *Department of Chemistry and Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China*

[2] *The Fifth People's Hospital of Shanghai, Fudan University, Shanghai 200240, China*

[3] *NHC Key Laboratory of Glycoconjugates Research (Fudan University), Shanghai, 200032, China*

[#] Equal contribution.

[*] Corresponding author(s).

E-mail: pyyang@fudan.edu.cn (Yang P), wqcao@fudan.edu.cn (Cao W)

[§]Current address:

Department of Chemistry and Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China

Running title: *Huang J et al / OGP: Association Repository of O-glycoproteins*

[a] ORCID 0000-0002-5316-4266

[b] ORCID 0000-0002-4780-8109

[c] ORCID 0000-0002-8856-266X

[d] ORCID 0000-0002-2829-2557

[e] ORCID 0000-0002-7971-9418

[f] ORCID 0000-0002-8579-0229

[g] ORCID 0000-0001-5779-1008

[h] ORCID 0000-0002-1007-4607

3125 words, 4 figures, 3 supplementary figures, 1 supplementary table, and 3 supplementary files; 29 References, 24 References in supplementary information and 16 References from 2014

Abstract

Numerous studies on cancer, biopharmaceuticals, and clinical trials have necessitated comprehensive and precise analysis of protein O-glycosylation. However, the lack of updated and convenient databases deters the storage and utilization of emerging O-glycoprotein data. To resolve this issue, an O-glycoprotein repository named OGP was established in this work. It was constructed with a collection of O-glycoprotein data from different sources. OGP contains 9354 O-glycosylation sites and 11,633 site-specific O-glycans mapping to 2133 O-glycoproteins, and it is the largest O-glycoprotein repository thus far. Based on the recorded O-glycosites, an O-glycosylation site prediction tool was developed. Moreover, an OGP-backed website is already available (*http://www.oglyp.org/*). The website comprises four specially designed and user-friendly modules: Statistic Analysis, Database Search, Site Prediction, and Data Submit. The first version of OGP repository and the website allow users to obtain vast O-glycoprotein related information, such as protein accession numbers, glycopeptides, site-specific glycan structures, experimental methods, and potential glycosylation sites. O-glycosylation data mining can be performed efficiently on this website, which can greatly facilitates O-glycosylation studies.

KEYWORDS: O-glycoproteins; Association repository; Site prediction;

Introduction

Comprehensive and precise analysis of O-glycoprotein would potentially further the current understanding of their roles in many physiological and pathological phenomena, such as intercellular communication [1], hereditary disorders, immune deficiencies, and cancer [2−4]. Great efforts have been made to analyse the complexity of O-glycosylation. And the recent technological advancements in many fields, especially mass spectrometry, generated impressive data on O-glycoproteins [5−11]. However, the lack of up-to-date and curated databases still hindered the archive, query, and utilization of emerging O-glycoprotein data.

Numerous studies have been performed to develop glycosylation-related databases [12−25]. However, most of these databases are focused on N-glycoproteins. Only a few databases contain data on O-glycoproteins. The most extensively used repository, UniCarbKB [13] provides massive N-glycoprotein data mainly and limited O-glycoproteins records. dbPTM [15, 16] is an integrated resource for post-translational modifications (PTMs) containing over 130 types of PTMs; however, it does not provide information regarding site-specific O-glycosylation. O-GLYCBASE [12] provides information regarding both glycans and glycosylation sites and is the most widely used database in O-glycosylation sttudies; however, it has not been updated since 2002. Moreover, it contains merely 189 O-glycoproteins and 2142 O-glycosylation sites, lagging in current O-glycoproteome data. In short, current O-glycoprotein databases are less satisfactory with notable issues including insufficient records, unknown data confidence, outdated or user-unfriendly (Supplementary Table 1).

It can be said that the dearth of O-glycoprotein databases has greatly impeded the development of O-glycosylation study. Recently, large-scale analysis of O-glycosylation sites and intact O-glycopeptides have gradually become possible. For example, Steentofh et al. developed a glycol-engineering method termed “SimpleCell” for large-scale identification of O-glycosylated sites [5]. Yang et al developed a method called “EXoO” for large-scale analysis of intact O-glycopeptides [7]. However, functional studies of O-glycoproteins are yet limited. In addition to the

complexity of O-glycosylation, another primary factor limiting studies on O-glycosylation is the difficulty in retrieving information from large data to select candidate O-glycoproteins. Thus, an updated O-glycosylation database providing curated information of protein O-glycosylation status, site-specific O-glycans, analytical methods, and other related information is largely in need and would accelerate studies on O-glycosylation.

In this study, an O-glycoprotein repository named OGP was constructed. OGP contains 9354 O-glycosylation sites and 11,633 site-specific O-glycans mapping to 2133 O-glycoproteins. To our knowledge, OGP is the most comprehensive experimentally characterized O-glycoprotein repository thus far. An O-glycosylation site prediction tool was also developed based on the recorded sites. An OGP-backed website is well established (*http://www.oglyp.org/*) to facilitate database access. The website contains four modules: Statistic Analysis, Database Search, Site Prediction, and Data Submit. All the aforementioned O-glycoprotein data can be easily obtained on the website. Such a comprehensive, user-friendly, and open-access O-glycoprotein repository would greatly benefit researches on O-glycosylation, development of O-glycoprotein drugs, or clinical studies.

Construction of the OGP repository

The OGP knowledgebase was constructed by integrating experimentally verified O-glycoproteins between 1998 and 2018 and other existing O-glycoprotein databases [12] (Fig. 1A). All proteins were manually curated, aligned with UniProt entries, and merged. Detailed extraction methods from the literature are described in Supplementary File S1. In total, 9354 glycosylation sites and 11,633 site-specific glycans mapping to 2133 glycoproteins of different species have been recorded in the database (Fig. 1B, C). The scale of the OGP repository is more than 20-fold bigger than the existing O-GlycBase 6.0, which is the largest O-glycoprotein repository thus far (Fig. 1D, E). This database is intended to be further updated to maintain its scale and content by expanding it with more published data.

The database records data such as proteins, peptide sequences, glycosylation sites, and site-specific glycans. For each site and site-specific glycan, detailed analytical strategies, such as sample sources, digestion enzymes, enrichment methods, analytical methods, and related studies are integrated. Besides, all glycoproteins recorded in the database have been aligned with their UniProt entries. Thus, additional data including protein sequence annotation, subcellular locations, and other post-translational modifications can be conveniently obtained. To better obtain topological information regarding glycans, a linear coding method (Supplementary File S2) has been used in this database to record site-specific glycan structures. Furthermore, experimental strategies for every glycopeptide, such as immunoprecipitation, gel filtration, and mass spectrometry, were manually extracted, verified, and recorded in the database. These data are easily retrievable from the OGP-backed website.

Development of an O-glycosylation site prediction model

Since O-glycosylation is highly complex but important, it is significant to better understand its modification pattern [26−29]. As a meaningful trial and for better utilization of OGP, an O-glycosylation site prediction model was developed using meticulously selected O-glycosylation sites with high confidence, recorded in OGP, which were identified at least one solid method to confirm its reliability and unambiguousness; for example, sites validated by Edman Sequencing or HCD and ETD tandem MS. The site prediction model was generated through three primary steps (Fig. 2A, Supplementary File S3): 1) construction of a dedicated training set; 2) optimization of parameters; 3) evaluation of site prediction performance. Through systematic optimization, a dedicated training dataset was established with a 1:1 positive: negative instances ratio for random sampling (1754 positive site-central sequences and 1754 negative site sequences) (Fig. 2B, Supplementary File S3-1.1). Sequences with 11 amino residues were considered preferable (Fig. 2C, Supplementary File S3-1.2). Thereafter, we compared the performance of different algorithms to predict O-glycosylation sites using Weka 3.8 as the data mining tool. The Random Forest algorithm displayed the best performance (Fig. 2D, E, Supplementary File S3-1.3). Finally, and was used to construct the prediction model.

A ten-fold cross validation was performed and results indicated that the prediction model has high accuracy and sensitivity (ROC value=0.982, Recall value=0.902) (Fig. 2D, E, Supplementary File S3-1.3).

Construction of the OGP-backed website

Based on the present OGP knowledge base, a dedicated website was constructed using hypertext markup language (HTML), cascading style sheet (CSS), and professional hypertext preprocessor(PHP); The design of the website is shown in Figure 3A. It contains three repositories in the underlying database layer: OGP, Prediction model, and Data submission. OGP is the core database that stores O-glycosylated protein sequences, sites, site-specific glycans, corresponding experimental data, and bibliography; The prediction model contained a model file and an inherent training set; Data submission is design to preserve user-uploaded information. By performing a set of actions including protein query and prediction model training and application in the operation layer, the website outputs four modules: Statistic analysis, Database search, Site prediction, and Data submit. The website is supported by most common web browsers such as Internet Explorer, Mozilla Firefox, Google Chrome, Safari and Opera.

Utility and the interface of the OGP website

The OGP-backed website, equipped with a user-friendly graphical interface, is already available (*http://www.oglyp.org/*) and comprises four main modules: Statistic analysis, Database search, Site Prediction, and Data Submit. Furthermore, other functions including database download, latest studies on O-glycosylation and useful links to databases (UniProt, UniCarbKB, and O-GlycBase) are provided and updated to provide more information for O-glycosylation study. The homepage of this website is shown in Figure 3B. Furthermore, the website provides detailed instructions and frequently asked questions (FAQ) to facilitate users.

The Statistic Analysis module provides an overview of the OGP repository, including the scale of total proteins, glycosylation sites, and site-specific glycans (Supplementary figure 1A), frequency distribution of O-glycoproteins in major taxonomic categories (Supplementary figure 1B), database scale comparison between

the OGP and the other primary databases (Supplementary figure 1C), data regarding glycoprotein-involved pathways, molecular and cellular functions of proteins, and related diseases determined through Ingenuity Pathway Analysis (Supplementary figure 1D, E, F). Furthermore, this module provides more essential information. For example, more than 95% of the reported O-glycoproteins are present in mammalians, 55% of which are present in *Homo sapiens*, indicating that O-glycosylation in other species warrants further analysis. All information would be updated in real time with the expansion of the OGP database.

In the Database Search module, users can retrieve O-glycoproteins flexibly by specifying the gene name, protein name, UniProt accession number, or glycan structure (Supplementary figure 2). Figure 4 shows a webpage returning results from a query of fibrinogen gamma chain as an example. These results comprise well-structured data on protein O-glycosylation including basic protein information (*i.e.*, protein name, UniProt accession number, and species, [Figure 4A]), a global map of all recorded O-glycosylation sites of the protein (Figure 4B), experimentally verified O-glycopeptides and site-specific glycans (Figure 4C), and corresponding experimental methods, identifiers and source references (Figure 4D, E).

The site prediction model developed herein has also been incorporated into the website to enable O-glycosylation site prediction. As is shown in Supplementary figure 3A, users can either fill out the template file with aligned site-centered sequences as instructed or simply upload a typical protein FASTA-type file and click on “predict” (Supplementary figure 3A). The prediction results for each site can be then displayed directly on the right side of the web page. Prediction scores range between 0 and 1; scores higher than 0.5 indicate positive sites, while those less than or equal to 0.5 indicate a highly probably non-O-glycosylated site. The higher the score, the greater the probability of a site being O-glycosylated and *vice versa.* The results can also be downloaded as shown in Supplementary figure 3B.

The Data Submit module enables users to upload new data into the OGP database and insert comments. All the new submitted data and comments are carefully recorded in a backend database and revised manually by experts at regular intervals.

Both a template form and an online form are accepted during a submission. What's more, when users upload the data by file, there will be a real time feedback shown below to inform users of those O-glycoproteins already in OGP database.

In addition, the database is accesible from OGP website. Download page can be found in the drop-down menu of tools on OGP homepage. The detailed top 500 entries could be directly downloaded (http://www.oglyp.org/download.php). Besides, there is a basic version of database, which provides all the O-glycoprotein Accessions and their corresponding O-glycosylation sites, for users to download freely (http://www.oglyp.org/download.php). The whole database could also be provided if users apply for it through E-mail request. The application guide is illustrated on the website (http://www.oglyp.org/download.php).

Conclusion and Discussion

The OGP repository, containing 9354 O-glycosylation sites and 11,633 site-specific O-glycans mapping to 2133 O-glycoproteins, is the most comprehensive, experimentally verified O-glycoprotein repository thus far. All data contained in the OGP repository have been manually curated, and protein sequences have been aligned with UniProt entries and merged. Based on recorded site data, an O-glycosylation site prediction tool has been developed to facilitate the prediction of O-glycosylation sites. The OGP-based website is already available (*http://www.oglyp.org/*) and contains four specially designed, user-friendly, functional modules: Statistic Analysis, Database Search, Site Prediction, and Data Submit. The initial version of the OGP repository and OGP-based website provides various data on O-glycoproteins, such as protein access number, sequence, site-specific glycan structures, experimental methods, and potential glycosylation sites. Mining of data on O-glycosylation can be carried out efficiently using this website. The OGP repository would greatly facilitate studies on O-glycosylation. The scale and content of this database is intended to be continuously expanded in subsequent versions of the OGP repository.

Availability

OGP prediction tool is freely available at http://www.oglyp.org/predict.php. OGP database is freely available at http://www.oglyp.org/download.php.

Authors' contributions

HJ developed the site prediction tool, participated in the database construction and helped to draft the manuscript. WM performed the database uniform, participated in the database and website construction and helped to draft the manuscript. ZY participated in website construction. KS participated in data collection. LM helped to revise the manuscript. JB participated in data collection. YP conceived the study, helped to draft the manuscript. CW designed the database, conceived the study, and draft the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors have declared no competing interests.

Acknowledgements

This work was supported by grants from the National Key Research and Development Program (2018YFC0910300, 2016YFA0501303 and 2016YFB0201702), the National Natural Science Foundation of China Project (91853102, 31600665 and 31570825). We would like to thank Mr. Jian-Qiang Wu to help with website construction, and Elsevier Premium Language Editing Services to help with the language editing.

References

[1] Hudson HF, Gerald WH, Ronald LS. Glycosylation precursors. In: Varki A, Richard DC, et al. editors. Essentials of Glycobiology, 3rd ed. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA; 2017,p.293–304.

[2] Boskovski MT, Yuan S, Pedersen NB, Goth CK, Makova S, Clausen H, et al. The heterotaxy gene GALNT11 glycosylates Notch to orchestrate cilia type and laterality. Nature 2013;504:456–9.

[3] Wolfert MA, Boons GJ. Adaptive immune activation: glycosylation does matter. Nat Chem Biol 2013;9:776–84.

[4] Campos D, Freitas D, Gomes J, Magalhães A, Steentoft C, Gomes C, et al. Probing the O-glycoproteome of gastric cancer cell lines for biomarker discovery. Mol Cell Proteomics 2015;14:1616–29.

[5] Steentoft C, Vakhrushev SY, Vester-Christensen MB, Schjoldager KTBG, Kong Y, Bennett EP, et al. Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. Nature Methods 2011;8:977–82.

[6] Ali L, Flowers SA, Jin C, Bennet EP, Ekwall AKH, Karlsson NG. The O-glycomap of lubricin, a novel mucin responsible for joint lubrication, identified by site-specific glycopeptide analysis. Mol Cell Proteomics 2014;13:3396–409.

[7] Yang W, Ao M, Hu Y, Li QK, Zhang H. Mapping the O-glycoproteome using site-specific extraction of O-linked glycopeptides (EXoO). Mol Syst Biol 2018;14:e8486.

[8] Zhang Y, Xie XF, Zhao XY, Tian F, Lv JC, Ying WT, et al. Systems analysis of singly and multiply O-glycosylated peptides in the human serum glycoproteome via EThcD and HCD mass spectrometry. J Proteomics 2018;170:14–27.

[9] Zhao XY, Zheng SS, Li YY, Huang JJ, Zhang WJ, Xie YP, et al. An Integrated Mass Spectroscopy Data Processing Strategy for Fast Identification, In-Depth, and Reproducible Quantification of Protein O-Glycosylation in a Large Cohort of Human Urine Samples. Anal Chem 2020;92:690–8.

[10] Mao JW, You X, Qin HQ, Wang CY, Wang LM, Ye, ML. A New Searching Strategy for the Identification of O-Linked Glycopeptides. Anal Chem 2019;91:3852–9.

[11] Qin HQ, Cheng K, Zhu J, Mao JW, Wang FJ, Dong MM, et al. Proteomics Analysis of O-GaINAc Glycosylation in Human Serum by an Integrated Strategy. Anal Chem 2017;89:1469–76.

[12] Gupta R, Birch H, Rapacki K, Brunak S, Hansen JE. O-GLYCBASE version 4.0: a revised database of O-glycosylated proteins. Nucleic Acids Res 1999;27:370–2.

[13] Campbell MP, Peterson R, Mariethoz J, Gasteiger E, Akune Y, Aoki-Kinoshita KF, et al. UniCarbKB: building a knowledge platform for glycoproteomics. Nucleic Acids Res 2014;42:D215–21.

[14] Malik A, Firoz A, Jha V, Ahmad S. PROCARB: A Database of Known and Modelled Carbohydrate-Binding Protein Structures with Sequence-Based Prediction Tools. Adv Bioinformatics 2010;2010:436036–9.

[15] Huang KY, Su MG, Kao HJ, Hsieh YC, Jhong JH, Cheng KH, et al. dbPTM 2016: 10-year anniversary of a resource for post-translational modification of proteins. Nucleic Acids Res 2016;44:D435–46.

[16] Huang KY, Lee TY, Kao HJ, Ma CT, Lee CC, Lin TH, et al. dbPTM in 2019: exploring disease association and cross-talk of post-translational modifications. Nucleic Acids Res 2019;47:D298–308.

[17] Bhat AH, Mondal H, Chauhan JS, Raghava GPS, Methi A, Rao A. ProGlycProt: a repository of experimentally characterized prokaryotic glycoproteins. Nucleic Acids Res 2012;40:D388–93.

[18] Cooper CA, Harrison MJ, Wilkins MR, Packer NH. GlycoSuiteDB: a new curated relational database of glycoprotein glycan structures and their biological sources. Nucleic Acids Res 2001;29:332–5.

[19] Cooper CA, Joshi HJ, Harrison MJ, Wilkins MR, Packer NH. GlycoSuiteDB: a curated relational database of glycoprotein glycan structures and their biological sources. Nucleic Acids Res 2003;31:511–3.

[20] Kao HJ, Huang CH, Bretaña NA, Lu CT, Huang KY, Weng SL, et al. A two-layered machine learning method to identify protein O-GlcNAcylation sites with O-GlcNAc transferase substrate motifs. BMC Bioinformatics 2015;16:S10.

[21] Kawasaki T, Nakao H, Takahashi E, Tominaga T. GlycoEpitope: the integrated database of carbohydrate antigens and antibodies. Trends Glycosci Glycotechnol 2006;18:267–72.

[22] Baycin-Hizal D, Tian Y, Akan I, Jacobson E, Clark D, Chu J, et al. GlycoFly: a database of Drosophila N-linked glycoproteins identified using SPEG--MS techniques. J Proteome Res 2011;10:2777–84.

[23] Zhang H, Loriaux P, Eng J, Campbell D, Keller A, Moss P, et al. UniPep--a database for human N-linked glycosites: a resource for biomarker discovery. Genome Biol 2006;7:R73.

[24] Kaji H, Saito H, Yamauchi Y, Shinkawa T, Taoka M, Hirabayashi J, et al. Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. Nat Biotechnol 2003;21:667–72.

[25] Kaji H, Yamauchi Y, Takahashi N, Isobe T. Mass spectrometric identification of N-linked glycopeptides using lectin-mediated affinity capture and glycosylation site-specific stable isotope tagging. Nature Protocols 2006;1:3019–27.

[26] Zhou K, Ai C, Dong P, Fan X, Yang L. A novel model to predict O-glycosylation sites using a highly unbalanced dataset. Glycoconj J 2012;29:551–64.

[27] Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J 2013;32:1478–88.

[28] Audagnotto M, Dal Peraro M. Protein post-translational modifications: In silico prediction tools and molecular modeling. Comput Struct Biotechnol J 2017;15:307–19.

[29] de Las Rivas M, Lira-Navarrete E, Gerken TA, Hurtado-Guerrero R. Polypeptide GalNAc-Ts: from redundancy to specificity. Curr Opin Struct Biol 2019;56:87–96.

Figure

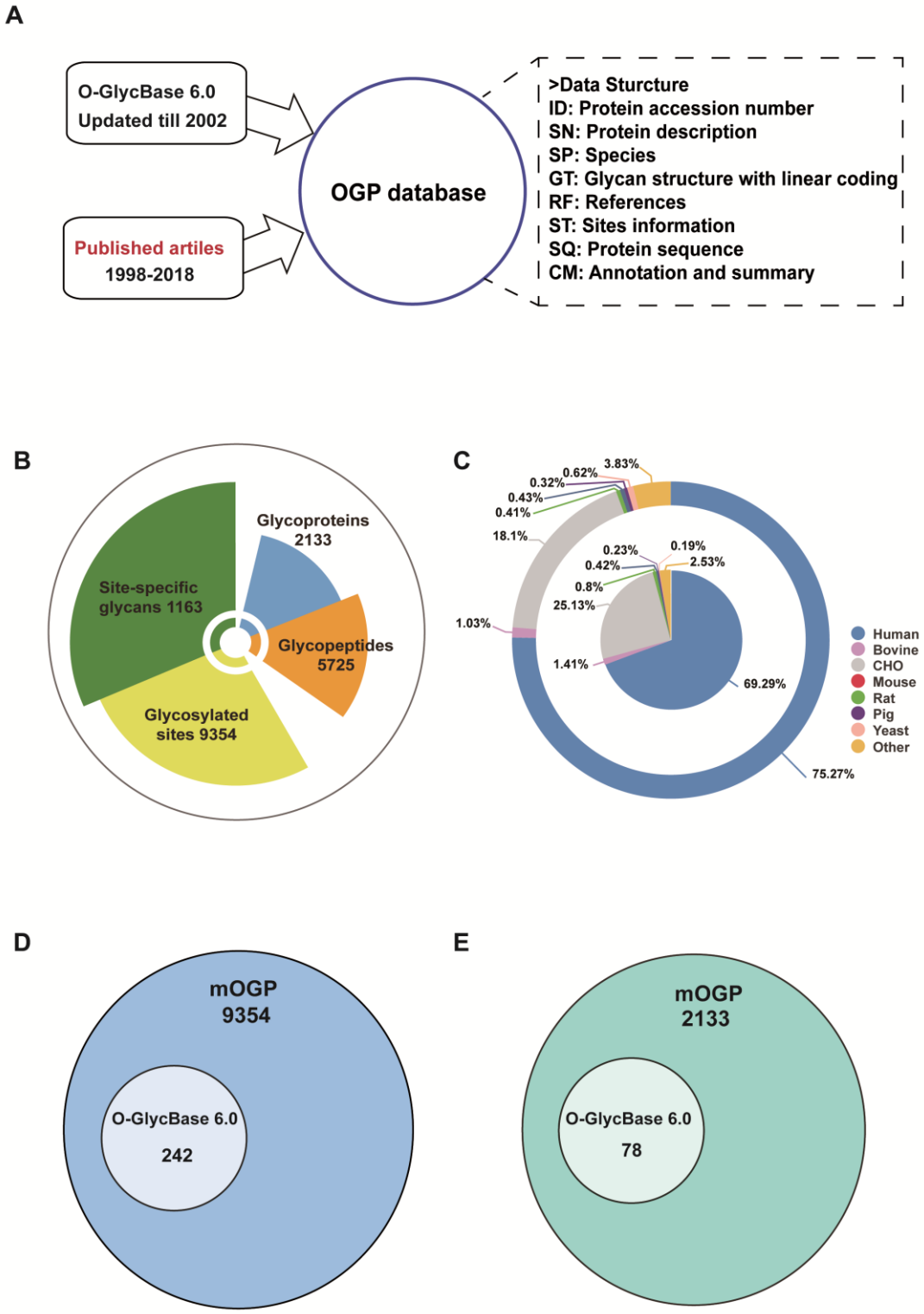

Figure 1 Overview of OGP repository.

A. OGP Data collection. **B.** The scale of the OGP repository. **C.** Species distribution of O-glycoproteins in OGP. **D.** Comparison of OGP with O-GlycBase 6.0 in

glycosylation site level. **E.** Comparison of OGP with O-GlycBase 6.0 in glycoprotein level.

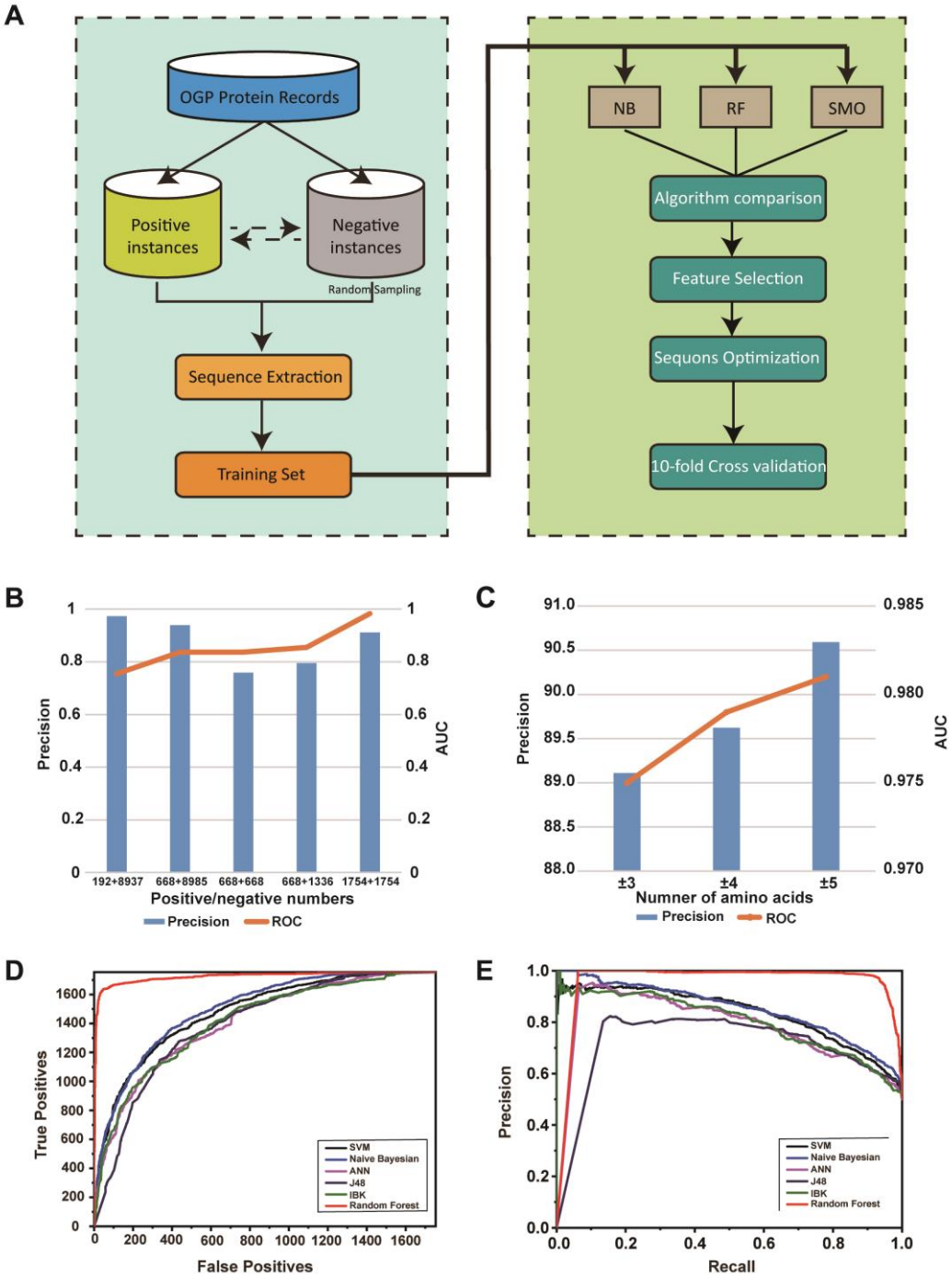

Figure 2 Development of O-glycosylation site prediction model

A. Workflow for building OGP based O-glycosite prediction model. **B.** Effect of positives and negatives scales and ratios on model prediction performance (AUC represented for area under curve). **C.** Influence of amino acid residuals on the

performance of site prediction model. **D.** ROC curves of each classification algorithm. **E.** PR curves of each classification algorithm.

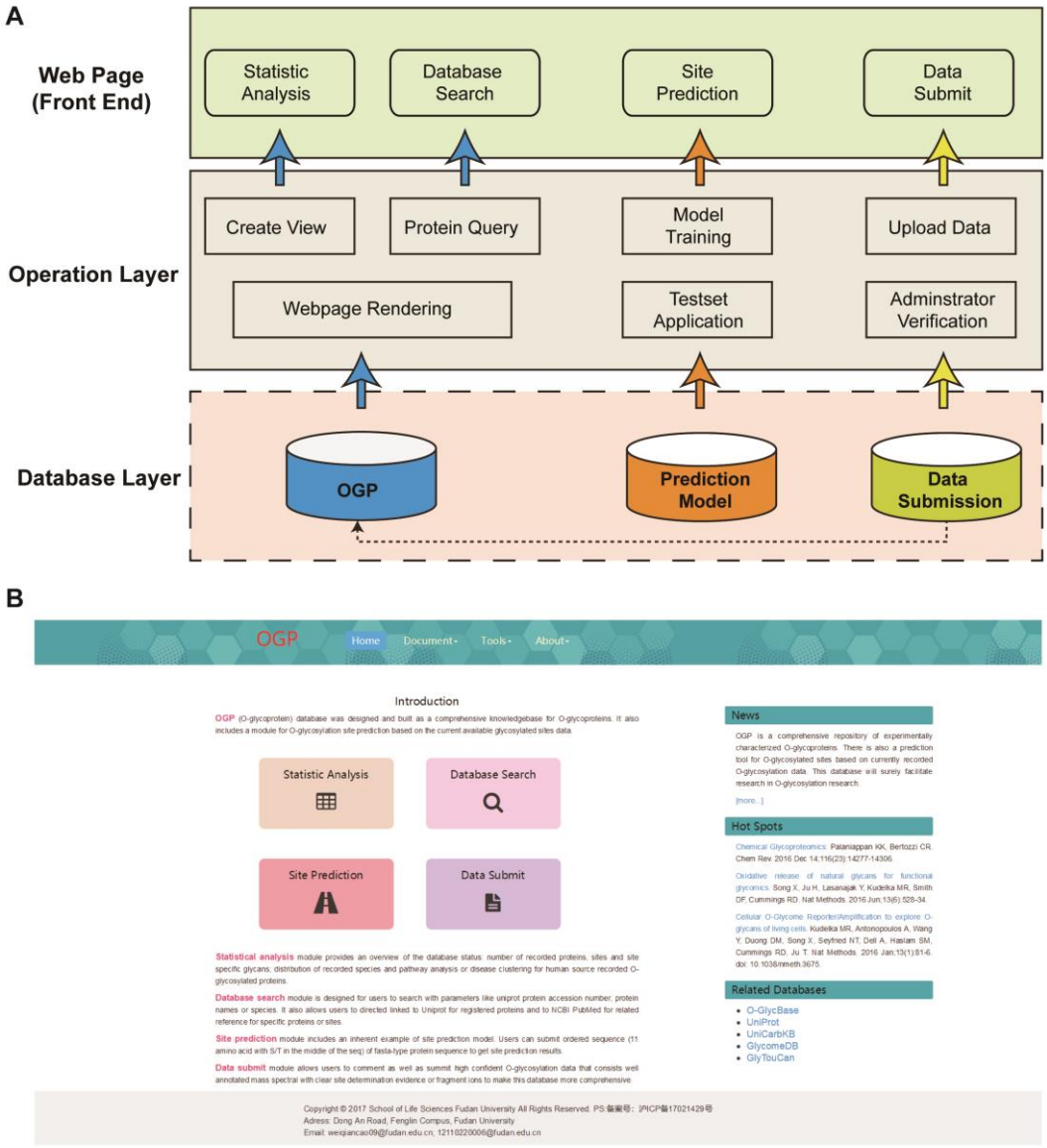

Figure 3 Construction of OGP-based website

A. The website structural design. **B.** Homepage of the website.

A

Result

Name

```
Accession   |  P02679 (click to Uniprot)
Isoforms    |  P02679-1
Protein     |  Fibrinogen gamma chain
Species     |  Homo sapiens (Human)
```

B

Protein Sequence

```
        10         20         30         40         50         60
MSWSLHPRNL ILYFYALLFL SSTCVAYVAT RDNCCILDER FGSYCPTTCG IADFLSTYQT
        70         80         90        100        110        120
KVDKDLQSLE DILHQVENKT SEVKQLIKAI QLTYNPDESS KPNMIDAATL KSRKMLEEIM
       130        140        150        160        170        180
KYEASILTHD SSIRYLQEIY NSNNQKIVNL KEKVAQLEAQ CQEPCKDTVQ IHDITGKDCQ
       190        200        210        220        230        240
DIANKGAKQS GLYFIKPLKA NQQFLVYCEI DGSGNGWTVF QKRLDGSVDF KKNWIQYKEG
       250        260        270        280        290        300
FGHLSPTGTT EFWLGNEKIH LISTQSAIPY ALRVELEDWN GRTSTADYAM FKVGPEADKY
       310        320        330        340        350        360
RLTYAYFAGG DAGDAFDGFD FGDDPSDKFF TSHNGMQFST WDNDNDKFEG NCAEQDGSGW
       370        380        390        400        410        420
WMNKCHAGHL NGVYYQGGTY SKASTPNGYD NGIIWATWKT RWYSMKKTTM KIIPFNRLTI
       430        440        450
GEGQQHHLGG AKQVRPEHPA ETEYDSLYPE DDL
```

C

Glycosylation Information

Peptide	Glyco Site	Glyco Struct	Reference Number
122YEASILTHDSSIR134	S132	HexNAc@S132	1
283TSTADYAMFK292	T285	HexNAc@T285	1
418LTIGEGQQHHLGGAK432	T419	HexNAc@T419	1
89AIQLTYNPDESSKPNMIDAATLK111	T93	HexNAc@T93	1
89AIQLTYNPDESSKPNMIDAATLK111	S99	HexNAc@S99	1
89AIQLTYNPDESSKPNMIDAATLK111	S100	HexNAc@S100	1
89AIQLTYNPDESSKPNMIDAATLK111	T109	HexNAc@T109	1

D

Method

Enezymes	Technique	Sample source	Enrich method	Reference Number
trypsin	LC-MSMS | ETD | HCD	HepG2ΔT1 and HepG2ΔT2 in WT SC	lectin	1

E

Reference

Number	Title	Pubmed Link	Journal
1	Deconstruction of O-glycosylation--GalNAc-T isoforms direct distinct subsets of the O-glycoproteome	EMBO Rep. 2015 Dec;16(12):1713-22.	EMBO Rep.

Figure 4 A webpage returning results from a query for Fibrinogen gamma chain

A. O-glycoprotein basic information. **B.** Protein Sequences and all recorded O-glycosylation sites with pink highlight. **C**. Experimentally verified O-glycopeptides and site-specific glycans. **D.** Corresponding experimental methods. D. Related source references.