GlycoPP V2.0 FRAMEWORK



Bioinformatics Centre, CSIR-Institute of Microbial Technology,

Chandigarh – 110036, India

Introduction
I. Navigation Panel4
1. Analyze Data4
2. Workflow5
3. Shared data library6
4 Help7
5. User
II. Tool Panel – GlycoPP2 tools7
1. Importing data to the GlycoPP27
2. Examples of GlycoPP V2.0 analysis8
3. GlycoPP V2.0 analysis19
III. Detail Panel
IV. History Panel

Introduction

GlycoPP V2.0 is a highly accurate glycosylation prediction made available for the analysis of prokaryotic protein sequences on the web based Galaxy Platform. GlycoPP prediction programmes are trained on the largest available and an extensive dataset of N- and O-linked glycosites extracted from experimentally characterized glycoproteins of prokaryotes as obtained from ProGlycProt V2.0(http://www.proglycprot.org/).

GlycoPP V2.0 is an enhanced and updated version of our GlycoPP V1.0 (http://crdd.osdd.net/raghava/glycopp). The workflow system is implemented using the open source workflow architecture, Galaxy. GlycoPP2 is freely available and can be accessed at https://ab-openlab.csir.res.in/alkarao/glycopp2/#

There are three modules in GlycoPP2, described in this manual. All the modules are accessible without registering to the system. However, for maintaining user-sessions it is recommended that anyone who is interested in creating data or task intensive workflow should register (**Figure**). The benefit of registration includes user sessions, saved histories, visualization, generation and execution of workflow and many others.

Ø	A we Link to the GlycoPP V2.0	GlycoP bserver for glycosite Link to the Shared Data of GlycoPP V2.0 galaxy	P V2.0 prediction in prokal Link to the workflow o Glycopp V2.0 galaxy	ryotes	Link to the Help Page GlycoPP V2.0
Home »	GlycoPP V2.0 Galaxy Platform »	Shared Data »	Example Workflow »	Team »	Help »
Overview of GlycoPP V2. GlycoPP V2.0 is a highly sequences on the web b an extensive dataset of 1 prokaryotes as obtained GlycoPP V2.0 is an enhal	accurate glycosylation prediction n ased Galaxy Platform. GlycoPP prec 4-glycosites and O-glycosites extra from ProGlycProt V2.0 (http://www nced and updated version of our G	nade available for the analysis of diction programmes are trained o cted from experimentally charact w.proglycprot.org/). Jyco DP U1.0 (http://crdd.ord.do. Send to the particular SVM N- and O-linked glycosylati and o-ample file in Glyco	prokaryotic protein n the largest available and erized glycoproteins of treachave/chropon model for the prediction of ion having preloaded fasta PP V2.0 galaxy platform	N-Inked Clycosylation	Grinked Gycosylation
r	O-linked Glycosylation SVM mod	el		N-linked Glycosylation SVM	nodel

Figure 1: GlycoPP V2.0 Main page

Galaxy GlycoPP V2.0	Analyze Data Workflow Visualize - Shared Data - Help - Login or Register	Using
	Log in or register a new account Welcome to Galaxy, please log in	
	Public name or Email Address	
	Password	
	Forgot password? Click here to reset your password.	
	Login	
	Don't have an account? Register here.	

Figure 2: GlycoPP2 Login Page

All the modules of GlycoPP2 and some default modules by Galaxy are accessible through a webbased interface which has following components (*Figure*).

Navigation Panel: It provides the links to major components of the server like Tools Page (Analyze Data), Workflow System, Shared Libraries, Visualization, Help Section and User Login/Registration.

Tool Panel: This panel lists all the tools available in GlycoPP2 along with default utilities in Galaxy.

Detail Panel (Canvas): This panel displays the interface of all the tools along with Input Parameters required to run a tool. It also provides help and examples to run a tool. This panel also displays the Output of a tool after its execution when user clicks on the eye sicon show in **History Panel**.

History Panel: This panel shows the information about the tools which are executed by a user. The information can include result after completion of a tool execution or error generated while running the tool. The workflow(s) are generated by extracting tasks from history panel.

Tools Panel	Navigation Panel	History I	Panel
Gali xy GlycoPP V2.0	Analyze Data Workflow Visualize * Shared Data * Help * User * 🚺		Using 191.4 KB
Tools 1/2 search tools Image: Constraint of the search tools Get Data Send Data	GlycoPP V2.0: A webserver for glycosite prediction in prokaryotes	i tory iearch datasets named history ipty)	2+0¢ 0
Lift-Over Expression Tools Collection Operations	Home » GlycoPP V2.0 Shared Data » GlycoPP V2.0 Example Workflow » GlycoPP V2.0 User Manual » Ann Cite Overview of GlycoPP V2.0	otation: k here to edit annotation	
Text Manipulation Convert Formats	GlycoPP is a webserver for predicting potential N-and O-glycos glycosite is an Asn residue and O-glycosite could be a serine or covalently and enzymatically at amide or hydroxyl group respect GlycoPP V2.0 is an updated version of our GlycoPP V1.0 (http://crdd.osdd.net/raghava/glycopp) developed for bit hydroxyl group respectively and the provide of the provided of the provided of the provided	This history is empty. Y your own data or get external source	/ou can load : data from an
Join, Subtract and Group Fetch Alignments/Sequences	riging accurate grycosylation prediction made available for the analysis of procarlyotic protein sequences on the web based Galaxy Platform. GlycoPP prediction programmes are trained on the largest available and an extensive dataset of N-glycosites and O-glycosites extracted from experimentally characterized glycoproteins of prokaryotes as obtained from ProGlycProt V2.0 (http://www.proglycprot.org/).		
Operate on Genomic Intervals Statistics Graph/Display Data	Data collection & Processing 41AA sequence pattern calculation Full length protein calculations		



I. Navigation Panel

1. Analyze Data

The data analysis page is where everything happens. There, you can run any available tools on the data, run complete workflows, browse or download a result, and share files with other users. It is

the default page when you open Galaxy in your browser, but you can also access it any time by clicking on "Analyze Data" in the Navigation Panel.

Galaxy GlycoPP V2.0	Analyze Data Workflow Visualize Shared Data Help User 🗰		Using 191.4 KB
Tools 🟠 🛓		History	2 + 🗆 🕈
search tools	G Analyze A webserver for glycosite prediction in	search datasets	8
Operate on Genomic Intervals	prokaryotes	Unnamed history	
Statistics		(empty)	۲
Graph/Display Data	Home » Glycopp V2.0 Shared Data » Glycopp V2.0 Example Workflow » Glycopp V2.0 User Manual »	Annotation:	
Phenotype Association		Click here to edit annotation	
GLYCOPP V2.0 ANALYSIS			
N-Linked Glycosylation	Glycosite is an Asn residue and O-glycosite could be a serine or threonine residue having a glycan attached	This history is empty your own data or get	: You can load et data from an
O-Linked Glycosylation	covalently and enzymatically at amide or hydroxyl group respectively. GivcoPP V2.0 is an updated version of our GivcoPP V1.0 (http://crdd.osdd.net/raphava/givcopp) developed for	external source	
EXAMPLES OF GLYCOPP V2.0 ANALYSIS	highly accurate glycosylation prediction made available for the analysis of prokaryotic protein sequences on the web based Galaxy Platform. GlycoPP prediction programmes are trained on the largest available and an		
Examples of N-Linked	extensive dataset of N-glycosites and O-glycosites extracted from experimentally characterized glycoproteins of prokaryotes as obtained from ProGlycProt V2.0 (http://www.proglycprot.org/).		
Glycosylation			
Examples of O-Linked	Pata collection 8. Processing		
Glycosylation	41AA sequence pattern calculation Full length protein calculations		
GlycoPP2 Tools	PreGlycProt SS SS PreGlycProt SS PreGlycProt SS SS PreGlycProt SS SS PreGlycProt SS PreGlycProtS		
Workflows	Removed 5-linked & DPC (Q, 0, 1, 2) PAAC (1, 1, 2) PSSM PPP		

2. Workflow

Workflows are analyses that are intended to be executed (one or more times) with different userprovided input Datasets. Workflow can be reused over and over, not only reducing tedious work, but enhancing reproducibility by applying the same exact methods to all of your data. Workflow is nothing but creating pipeline, user can use it again and again or user can published it.

Workflow can be created through navigation panel or from tool panel. In workflow section user can create workflow or can upload or import the workflow. The canvas is where inputs, tools, and noodles are added and connected as you build and modify your workflow (Figure 4). Selecting Edit opens the workflow editor view (Figure 3). The navigator provides a full view of your workflow in a condensed format (Figure 4). Accessed by clicking on the gear icon on the right side of the center Workflow Canvas upper bar, the workflow editor menu (Figure 3) is for global editor actions. It consists of Save, Run, Edit Attributes, Auto re-layout, Close

1. Workflow	
Galaxy GlycoPP V2.0 Analyze Data Workflow Visualize • Shared Data • Help • User •	Using 237.4 KB
Tools Import search tools Search tools Get Data Search name, annotation, owner, Q Advanced Search 2. Saved Workflow Advanced Search 2. Saved Workflow Lift-Over Expression Tools Collection Operations 32 minutes Text Manipulation Operation Convert Formats Filter and Sort Join, Subtract and Group Fetch Alignments/Sequences	History 2 + 1 2 search datasets 2 Unnamed history (empty) This history is empty. You can load your own data or get data from an external source

Figure 5: Options for workflow

The following example of workflow shows the "Prediction of N- and O-linked glycosylation prediction". The prediction workflow can use the four implemented svm model for N-linked glycosyation and six svm model for O-linked glycosylation

> Run Workflow for N- and O-linked glycosylation \

Galaxy	SIYCOPP V2.0	Analyze Data Workflow Visualize * Shared Data * Help * Use			Using 264.6 KB
Tools	☆ 🛓			History	2 + 🗆 🕈
search tools	0	Workflow: Workflow for O-Linked Glycosylation predictions		search datasets	8
Operate on Geno			 Run workflow 	Unnamed history	
Statistics Graph/Display Da	ata	History Options		(empty)	>
Phenotype Associ	iation YSIS	Send results to a new history Yes No		This history is emp your own data or	oty. You can load get data from an
N-Linked Glycosy	/lation	1: Input dataset	۲	external source	
O-Linked Glycosy EXAMPLES OF GLYCO ANALYSIS	vlation	D 2 No fasta, fasta, fasta, fasta, fasta or fasta dataset available.	• 🕞		
Examples of N-Lir	nked		۲		

3. Shared data library

Data libraries are collections of Datasets that are accessible from within a Galaxy instance. Libraries are designed for sharing datasets in between users or groups. The data library of GlycoPP2 consists of prokaryotic glycoproteins list protein list used in SVM model generation. Some of the actions that can be performed on data libraries are accessed by clicking the pop-up menu icon just right of the data library name.

- View Information –Shows the information about dataset.
- Import this dataset into your current history this creates an item in your current

Figure 6: Workflow overview

history on which you can perform analysis. The item is a pointer to the library dataset disk file, so the file is not copied on disk.

• Download this dataset - this allows you to download a local copy of the dataset.

1. Shared Data library

Galaxy GlycoPP V2.0	Analyze Data Workflow Visualize 🔻	′Shared Data ▼ Help ▼ User ▼ 🗱	Using 191.4 KB
Search			
Name \downarrow_2^A	Description	Synopsis	
Example fasta files	Single fasta and multifasta files for as	Tools and workflow input fasta files	0
GlycoPP V2.0 Testing dataset	Testing dataset	Dataset used in testing SVM model	0
GlycoPP V2.0 Training dataset	Training dataset	Dataset used in training SVM model	0
Glycopp2 bpp dataset	dataset for tool testing		0
	« 0 1 2 »	20 per page, 4 total	

Figure 6: Viewed for shared library

4 Help

The help section of Galaxy consists of Support, Search, Mailing List, Videos, Wiki and How to cite Galaxy. User can fine user manual in help section.

5. User

Login option and register option can get in user section. It is recommended that user register their account before using framework. Although unregistered users have access to tools available but their history is stored temporarily. On the other hand, registered users can save and retrieve their results in history panel later too.

Tool Panel – GlycoPP2 tools

The user can get tools in tool panel. There are two categories of tools galaxy inbuilt tool and GlycoPP2 tools. Galaxy tools consists of Data importing, Manipulation, Filtering, Sorting, Format conversion etc. GlycoPP2 tool are specific for finding N- and O-linked glycosite in prokaryotic protein sequence

1. Importing data to the GlycoPP2

A user can upload the data using the Galaxy tool **Get Data.** The uploaded data can be used for GlycoPP V2.0 Analysis. The following figure shows the file upload method.

1. Get data / Upload file

Salaxy GlycoPP	V2.0 Analyze Data Workflow Visualize * Shared Data * Heln * User *	Using 191.4 KB
Tools 🗘	Download from web or upload from disk	y C+⊞¢
search tools	Regular Composite Collection Rule-based	h datasets
Statistics	You added 1 file(s) to the queue. Add more files or click 'Start' to proceed.	ned history
Graph/Display Data	Name Size Type Genome Settings Status	••
Phenotype Association	🗭 New File - Auto-de V Q Additional S V 🗘 🕫	s history is empty. You can load
GLYCOPP V2.0 ANALYSIS	Download data from the web by entering URLs (one per line) or directly paste content.	ir own data or get data from an
N-Linked Glycosylation		
O-Linked Glycosylation		
EXAMPLES OF GLYCOPP V2.0 ANALYSIS		
Examples of N-Linked Glycosylation		
Examples of O-Linked Glycosylation	Type (set all): Auto-detect Q Genome (set all): Additional S V	
GlycoPP2 Tools	Chonce Incel file Pacte/Eatch data Dause Reset Start Close	
Workflows		ve te Windows

Figure 8: Get Data screen

2. Examples of Glycopp V2.0 Analysis

These have example fasta and multifasta file reloded for the prediction analysis of Nand O-linked glycosylation SVM model. BPP, BPP+ASA, BPP+SS and BPP+ASA+SS for N-linked Glycosylation and CTD, PAAC, SER, CPP+SS, DPC+SS and DPC+ASA for Olinked Glycosylation

GlycoPP V2.0 A webserver for glycosite prediction in prokaryotes	CARAY CIVCLEP V2.0
Nume Operation Nume Num Nume Nume	I winder in the second of t
Send to the particular SVM mod linked glycosylation having p GlycoPP V2.0	lel for the prediction of N- and O- preloaded fasta example file in galaxy platform

Figure 9: Example of GlycoPP V2.0 Analsis Screen

a) Example of N-linked Glycosylation

BPP example Binary Profile Pattern based prediction example



Figure-10 : Example of GlycoPP V2.0 Analsis overview

Galaxy GlycoPP V2	.0	Analyze Da	ta Workflow	Visualize 👻 Shared Data '	 Help ▼ Login or Reg 	ister		Using 3.5 KB
Tools 🔔		>Q0PAM0		Length = 171		^	History	C 🕈
search tools		>Q0PAM0 Potential N-Linked Glycosyl	ated Sites:	Length = 171			search datasets	0
Fetch Alignments/Sequences	*	MAIKIFGILIALFTITFTILSLQD	PYSLNLQTNALN	KNIEAKNLKAYESNTSIIKAY	YKANSWVRYADRDEFNDF	ITLNLDFNLSANRLEFFNKDM	Unnamed history	
Operate on Genomic Intervals		LFEGNVTYIGSNNVKIISQEVE	YQPKDKILHTNTI	VFKALINGSIINGNTLNYDV	(NKILNIQGVNAWLQDK	_	1 shown, 1 deleted	
Statistics	ь	GlycoPP v2.0 Prediction Me	thod = BPP	SVM Threshold =	0		3.47 KB	
Graph/Display Data	11	Position		Residue	Score	Prediction	2: BPP example	● # ×
Phenotype Association	11			NIO	0.68603250	Non alwandated		
GLYCOPP V2.0 ANALYSIS	11	33		NAL	-0.72174881	Non-glycosylated		
N-Linked Glycosylation	Ľ	36		NFK	-0.12207736	Non-glycosylated		
O Linked Chrosselation		39		NIE	-0.32470975	Non-glycosylated		
O-Linked Glycosylation		44		NLK	-0.20398314	Non-glycosylated		
EXAMPLES OF GLYCOPP V2.0		51		NTS	1.0002212	Potential Glycosylated		
ANALYSIS		62		NSW	-0.52280092	Non-glycosylated		
Examples of N-Linked		74		NDF	-0.35368478	Non-glycosylated		
Glycosylation		80		NLD	-0.8578013	Non-glycosylated		
RDD example Ripper/ Profile Pattern		84		NLS	0.71732003	Potential Glycosylated		
based prediction example		88		NRL	-0.36008792	Non-glycosylated		
RPP+ASA example (Binany Profile		94		NKD	-0.15063364	Non-glycosylated	Activate Windows	
Composition + Accessible Surface		105		NVT	0.63372566	Potential Glycosylated	Go to Settings to activat	
Aras) based prediction example	*	112		NNV	-0.70692049	Non-glycosylated 👻		

Figure 11: Example of GlycoPP V2.0 Analsis output for BPP (Binary Profile Pattern) forN-linked glycosylation

BPP+ASA example (Binary Profile Composition+ Accessible Surface Area) based prediction example



Figure 12: Example of GlycoPP V2.0 Analsis overview

Galaxy GlycoPP V2.0	Analyze Data Wo	orkflow Visualize - Shared Data -	Help - Login or Regis	ster		Using 3.5 KB
Tools 🔔	>Q0PAM0	Length = 171		^	History	2 \$
search tools	>Q0PAM0	Length = 171			search datasets	0
Fetch Alignments/Sequences	Potential N-Linked Glycosylated Si MAIKIFGILIALFTITFTILSLQDPYSLNL	tes: .QTNALNFKNIEAKNLKAYESNTSIIKAYYK	ANSWVRYADRDEFNDFIT	LNLDFNLSANRLEFFNKDM	Unnamed history	
Operate on Genomic Intervals	LFEGNVTYIGSNNVKIISQEVEYQPKE	- DKILHTNTNFKALINGSIINGNTLNYDVKN	KILNIQGVNAWLQDK		1 shown, 1 deleted	
Statistics	GlycoPP v2.0 Prediction Method =	BPP SVM Threshold = 0			3.47 KB	
Graph/Display Data	Position	Residue	Score	Prediction	2: BPP example	• / ×
Phenotype Association		NIO	0.68603350	Nep always dated		
GLYCOPP V2.0 ANALYSIS	33	NAL	-0.72174881	Non-glycosylated		
N-Linked Glycosylation	36	NFK	-0.12207736	Non-glycosylated		
O Linked Glycosylation	39	NIE	-0.32470975	Non-glycosylated		
O-Linked Glycosylation	44	NLK	-0.20398314	Non-glycosylated		
EXAMPLES OF GLYCOPP V2.0	51	NTS	1.0002212	Potential Glycosylated		
ANALTSIS	62	NSW	-0.52280092	Non-glycosylated		
Examples of N-Linked	74 90	NUF	-0.35368478	Non-glycosylated		
Glycosylation	84	NUS	0.71732003	Potential Glycosylated		
BPP example Binary Profile Pattern	88	NRL	-0.36008792	Non-glycosylated		
PDD I ASA example (Pipapi Profile	94	NKD	-0.15063364	Non-glycosylated	Activata Mindowa	
Composition + Accessible Surface	105	NVT	0.63372566	Potential Glycosylated	Go to Settings to activate	
Area) based prediction example	112	NNV	-0.70692049	Non-glycosylated 👻		

Figure 13: Example of GlycoPP V2.0 Analsis output for BPP+ASA (Binary Profile Pattern+ Accessible Surface Are) forN-linked glycosylation

BPP+ASA+SS example (Binary Profile Composition+ Accessible Surface Area + Secondary Structure) based prediction example



Figure 14: Example of GlycoPP V2.0 Analsis overview

Galaxy GlycoPP	V2.0	Ar	alyze Data	Workflow Visualiz	e ▼ Shared Data ▼ H	lelp 👻 Login or Register			Using 3.5 KB
Tools	±	>Q0PAM0			Length = 171			History	C 🕈
search tools	8	>Q0PAM0			Length = 171			search datasets	8
Get Data	^	Potential N-Linked MAIKIFGILIALFTITF	Glycosylate TILSLQDPYS	Unnamed history					
Send Data		LFEGNVTYIGSNNV	KIISQEVEYQ	PKDKILHTNTNFKALII	NGSIINGNTLNYDVKNKIL	NIQGVNAWLQDK		2 shown	
Lift-Over	- 1	GlycoPP v2.0 Predi	ction Metho	d = BPP+ASA+SS	SVM Threshold =	= 0		3.5 KB	
Expression Tools		Position			Residue	Score	Prediction	2: BPP+SS example	• / ×
Collection Operations	- 1	29			NLQ	-0.61250663	Non-glycosylated	1: BPP+ASA+SS example	• / ×
Text Manipulation		33			NAL	-0.53444309	Non-glycosylated		
Convert Formats		36			NFK	0.026839411	Potential Glycosylat		
		39			NIE	-0.29871205	Non-glycosylated		
Filter and Sort		44			NLK	-0.30010003	Non-glycosylated		
Join, Subtract and Group		51			NTS	0.58642466	Potential Glycosylat		
Fetch Alianments/Sequence	c	62			NSW	-0.39528659	Non-glycosylated		
Teten Alignments/Sequence	3	74			NDF	-0.20714487	Non-glycosylated		
Operate on Genomic Interva	als	80			NLD	-0.65906657	Non-glycosylated		
Statistics		84			NLS	0.55723666	Potential Glycosylat		
Cranh (Dianlas) Data		88			NRL	-0.28949737	Non-glycosylated		
Graph/Display Data		94			NKD	-0.23029647	Non-alvcosvlated		

Figure 15: Example of GlycoPP V2.0 Analsis output for BPP+ASA+SS example (Binary Profile Composition+ Accessible Surface Area + Secondary Structure) forN-linked glycosylation

BPP+SS example (Binary Profile Composition+ Secondary Structure) based prediction example



Figure 16: Example of GlycoPP V2.0 Analsis overview

Tools 🛓	>Q0PAM0	Length = 171		A	History	C 🕈					
search tools	>Q0PAM0 Potential N-Linked Glycosylated Sites:	Length = 171			search datasets	0					
Operate on Genomic intervais	MAIKIFGILIALFTITFTILSLQDPYSLNLQTNALNFKN	taikifgilialFtitFtilslqDpyslnlqtnalnFknieaknlkayesntsiikayykanswvryadrdeFndFitlnldFnlsanrleFFnkdv									
Statistics	LFEGNVTYIGSNNVKIISQEVEYQPKDKILHTNTNFK	alingsiingntlnydvkn	IKILNIQGVNAWLQDK		2 shown						
Graph/Display Data	GlycoPP v2.0 Prediction Method = BPP+SS	SVM Threshold =	: 0		3.5 KB						
Phenotype Association	Position	Residue	Score	Prediction	2. BPD+SS example						
GLYCOPP V2.0 ANALYSIS					L. Di ti i bo ciumpic						
N-Linked Glycosylation	29	NLQ	-0.61294694	Non-glycosylated	1: BPP+ASA+SS example	👁 🖋 🗙					
O Linked Chronylation	33	NAL	-0.53504833	Non-glycosylated							
O-Linked Glycosylation	36	NFK	0.026195499	Potential Glycosylate							
EXAMPLES OF GLYCOPP V2.0	39	NIE	-0.29947346	Non-glycosylated							
ANALYSIS	44	NLK	-0.300722	Non-glycosylated							
Examples of N-Linked	51	NTS	0.58505686	Potential Glycosylate							
Glycosylation	62	NSW	-0.39570128	Non-glycosylated							
Even and a fight d	74	NDF	-0.20760915	Non-glycosylated							
Examples of O-Linked	80	NLD	-0.65981998	Non-glycosylated							
Glycosylation	84	NLS	0.55655757	Potential Glycosylate							
GlycoPP2 Tools	88	NRL	-0.28977783	Non-glycosylated							

Figure 17: Example of GlycoPP V2.0 Analsis output for BPP+ASA+SS example (Binary Profile Composition+ Accessible Surface Area + Secondary Structure) forN-linked glycosylation

b) Example of O-linked Glycosylation

CTD_example Conjoint Triad Descriptors Example



Figure 18: Example of GlycoPP V2.0 Analsis overview

6 Galaxy Glyc	oPP V2.0		Analyze Data	Workflow	Visualize 🔻	Shared Data	r Help ▼	Login or Register			Using 5.8 KB
Tools	<u>*</u>	>B4EB72			Le	ength = 171			^	History	C 🕈
search tools	8	>B4EB72		1.01	Le	ength = 171				search datasets	8
operate on Genomic in	ונכו עמוס	Potential O-Li	NKED GIYCOSYIAT	Ed Sites:	ISCROSPREAT		CHITHING		ESCTOTTRVNAHDDS	Unnamed history	
Statistics		DSTPPDVNGF	GISLKIGSVDYQ	1 shown, 2 deleted							
Graph/Display Data		GlycoPP v2.0	Prediction Meth	od = CTD	S	/M Threshold =	0			5.8 KB	
Phenotype Association		Position			Re	esidue		Score	Prediction	D. CTD annuals	a b u
GLYCOPP V2.0 ANALYSIS										5: CTD_example	• / *
N-Linked Glycosylation	n	2			Т			-0.20935397	Non-glycosylated		
O Linked Chronylation		7			Т			-0.19029675	Non-glycosylated		
O-Linked Glycosylation		15			Т			-0.1508115	Non-glycosylated		
EXAMPLES OF GLYCOPP V2	.0	20			S			-0.40016528	Non-glycosylated		
ANALYSIS		25			Т			-0.11553263	Non-glycosylated		
Examples of N-Linked		30			S			-0.13035721	Non-glycosylated		
Glycosylation		33			S			0.039000312	Potential Glycosylat		
		37			S			-0.11382714	Non-glycosylated		
Examples of O-Linked		40			S			-0.11227607	Non-glycosylated		
Glycosylation		43			S			-0.080158341	Non-glycosylated		
GlycoPP2 Tools		49			т			0.030458067	Potential Glycosylat		
		50			-			0.400006.44	D 1 1 1 Cl 1 1		

Figure 19: Example of GlycoPP V2.0 Analsis output for CTD (Conjoint Triad Distributor) for O-linked glycosylation

PAAC_example Pseudo Amino Acid Composition based prediction example



Figure 21: Example of GlycoPP V2.0 Analsis overview

Galaxy GlycoPP V2	2.0	Analyze Data Workflow	v Visualize 🛪 S	hared Data 👻 Help 🔻	Login or Registe	r 💶		Using 8.1 KB
Tools 👲	>B4EB72		Leng	jth = 171		<u>^</u>	History	C 🕈
search tools	>B4EB72	aked Glycosylated Sites:	Leng	th = 171			search datasets	8
CTD_example Conjoint Triad Descriptors Example	MTDPRHTVRI DSTPPDVNGF GhrcePR v2.0	AVGATALGVSALGATLPACSA GISLKIGSVDYQMPYQPVQS	Unnamed history 1 shown, 3 deleted					
PAAC_example Pseudo Amino Acid Composition based prediction example	Position	rediction Method = PAAC	Resid	due	Score	Prediction	8.09 KB 4: PAAC_example	⊻ ⊛ & ×
SER example Shannon Entropy of Residues based prediction example	2		T		-0.20271261	Non-glycosylated		
CPP+SS example (Composition Profile of Patterns + Secondary	7 15		T		-0.18153102 -0.047441369	Non-glycosylated Non-glycosylated		
DPC+SS example (Dipeptide Composition + Secondary Structure)	25 30		T S		-0.15297738 -0.18510155	Non-glycosylated Non-glycosylated Non-glycosylated		
based prediction example DPC+ASA example (Dipeptide Composition + Accessible Surface	33 37		s s		-0.18575067 -0.11436738	Non-glycosylated Non-glycosylated		
Area) based prediction example GlycoPP2 Tools	40 43 49		S S T		-0.082389054 -0.010831194	Non-glycosylated Non-glycosylated		

Figure 22: Example of GlycoPP V2.0 Analsis output for PAAC (Pseudo Amino Acid Composition) for O-linked glycosylation

SER example Shannon Entropy of Residues based prediction example



Figure 23: Example of GlycoPP V2.0 Analsis overview

Galaxy GlycoPP \	/2.0	Analyze Data Wo	rkflow Visualize - Shar	ed Data ▼ Help ▼ Login o	r Register		Using 10.4 k
Tools	<u>★</u> >B4EB72		Length	= 171		History	21
search tools	>B4EB72		Length	= 171		cearch datasets	0
search tools	Potential O	Linked Glycosylated Sit	es:				
Get Data	MTDPRHTV	RIAVGATALGVSALGATLF	ACSAHSGPGSPPSAPSAPA	AATVMVEGHTHTISGVVECRTS	PAVRTATPSESGTQTTRVNAHDDS	Unnamed history	
Send Data	DSTPPDVN	GFGISLKIGSVDYQMPYQ	PVQSPTQVEATRQGKSYTLT	GTGHAVIPGQTGMRELPFGVH	/TCP	1 shown, 4 deleted	
	GlycoPP v2	0 Prediction Method =	SER SVM Th	reshold = 0		10.38 KB	
Lift-Over							
Expression Tools	Position		Residue	Score	Prediction	5: SER example	۰ 🖋 ک
Collection Operations	2		т	-0.36043	891 Non-glycosylated		
Text Manipulation	7		Т	-0.53629	17 Non-glycosylated		
Convert Formats	15		Т	-0.29736	866 Non-glycosylated		
	20		S	-0.15835	57 Non-glycosylated		
Filter and Sort	25		Т	-0.18603	931 Non-glycosylated		
Join, Subtract and Group	30		S	-0.13741	532 Non-glycosylated		
Fetch Alignments/Sequences	33		S	-0.16468	103 Non-glycosylated		
reten Algiments/Sequences	37		S	-0.10327	306 Non-glycosylated		
Operate on Genomic Intervals	40		S	0.169394	45 Potential Glycosyla		
Statistics	43		S	0.04404	Potential Glycosyla	16	
Graph (Display Data	49		Т	0.06638	Potential Glycosyla	ie	

Figure 24: Example of GlycoPP V2.0 Analsis output for SER (Shannon Entropy of Residues) for O-linked glycosylation

CPP+SS example (Composition Profile of Patterns + Secondary Structure) based prediction example

Galaxy GlycoPP V	2.0 Analyze Data Workflow Visualize * Shared Data * Help * Login or Register	Using 0 bytes
Tools 1 search tools	CPP+SS_example (Composition Profile of Patterns + Secondary Structure) based prediction example (Galaxy Version 0.1.1)	History 2 ¢ search datasets
based prediction example	Preloaded example for input fasta sequence	Unnamed history
Examples of O-Linked Glycosylation	> 84E872 MTDPRHTVRIAVGATALGVSALGATLPACSAHSGPGSPPSAPSAPAAATVMVEGHTHTISGVVECRTSPAVRTATPSESGTQTTRVNAHDDSASVTLSLS	(empty)
CTD_example Conjoint Triad Descriptors Example	DSTPPDVNGFGISLKIGSVDYQMPYQPPQSPTQVEATRQGKSYTLTGTGHAVIPGQTGMRELPFGVHVTCP	This history is empty. You can load
PAAC_example Pseudo Amino Acid Composition based prediction example	Input fasta sequence (Allowed special character in fasta header line ' - = () / + * ^ . ? ? []'. Invalid character will replaced by 'X') Select SVM Threshold:	external source
SER_example Shannon Entropy of Residues based prediction example	•	
CPP+SS_example (Composition Profile of Patterns + Secondary Structure) based prediction example	SVM Threshold value between -1.00 to +1.00. Default is 0 ✓ Execute	
DPC+SS_example (Dipeptide Composition + Secondary Structure) based prediction example	INFO: CPP+SS (Composition Profile of Patterns + Secondary Structure) based prediction :	
DPC+ASA_example (Dipeptide Composition + Accessible Surface Area) based prediction example	Composition promie of patterns is the percentage requencies of each amino acid in a trixed length sequence patterns. Previous studies on eukaryotic glycoproteins suggested that the probability of finding glycosite was higher at positions where there was a secondary structure change. So, we applied the secondary structure along with DPC combination to get the higher result.	
CI DDO T I	Please provide a fasta file for GlycoPP V2.0 SER prediction tool:-	

Figure 25: Example of GlycoPP V2.0 Analsis overview

🕤 Galaxy GlycoPP	V2.0	An	alyze Data N	Workflow '	Visualize 🔻	Shared Data 🔻	Help 🕶	Login or Register			Using 12.7 KB
Tools	1	>B4EB72				Length = 171			-	History	C 🕈
search tools	8	>B4EB72	Cl	search datasets	8						
Ciycosynation		Potential O-Linked	Giycosylated	Unnamed history							
CTD_example Conjoint Triad Descriptors Example		DSTPPDVNGFGISL	(IGSVDYQMP	1 shown, 5 deleted							
PAAC_example Pseudo Amino Aci	d	GlycoPP v2.0 Prediction Method = CPP+SS				SVM Threshold =	= 0		12.68 KB		
example		Position				Residue		Score	Prediction	6: CPP+SS example	• # ×
SER example Shannon Entropy of Residues based prediction example	2	2				т		-0.29620787	Non-glycosylated		
CPR+SS example (Composition		7				Т		-0.52232343	Non-glycosylated		
Profile of Patterns + Secondary		15				Т		-0.39875222	Non-glycosylated		
Structure) based prediction example	e	20				S		-0.19466643	Non-glycosylated		
DPC+SS example (Dipeptide		25				Т		-0.20355421	Non-glycosylated		
Composition + Secondary Structur	e)	30				S		-0.20582699	Non-glycosylated		
based prediction example		33				S		-0.20324	Non-glycosylated		
DPC+ASA example (Dipeptide		37				S		-0.10289745	Non-glycosylated		
Composition + Accessible Surface		40				S		0.068873432	Potential Glycosyla		
Area) based prediction example		43				S		-0.051108665	Non-glycosylated		
GlycoPP2 Tools		49				Т		0.0087108772	Potential Glycosyla		
						-					

Figure 26: Example of GlycoPP V2.0 Analsis output for CPP+SS (Composition Profile of Patterns + Secondary Structure) for O-linked glycosylation

DPC+SS example (Dipeptide Composition + Secondary Structure)based prediction example

Galaxy GlycoPP V	2.0 Analyze Data Workflow Visualize * Shared Data * Help * Login or Register	Using 0 bytes										
Tools .	DPC+SS_example (Dipeptide Composition + Secondary Structure) based prediction example (Galaxy Version 0.1.1)	History 2 ¢										
based prediction example	Preloaded example for input fasta sequence	Unnamed history										
Examples of O-Linked Glycosylation	> B4E872 MTDPRHTVRIAVGATALGVSALGATLPACSAHSGPGSPPSAPSAPAAATVMVEGHTHTISGVVECRTSPAVRTATPSESGTQTTRVNAHDDSASVTLSLS DSTDPRIVSCHUNGCHUNGALDNOR VSS DTAVE ADDREGENT TSTSLLNINGCOTA NOL DSCH USED	(empty)										
CTD_example Conjoint Triad Descriptors Example		STPPDVNGFGISLKIGSVDVQMPYQPVQSPTQVEATRQGKSYTLTGTGHAVIPGQTGMRELPFGVHVTCP										
PAAC_example Pseudo Amino Acid Composition based prediction example	Input fasta sequence (Allowed special character in fasta header line ' - = () / + * ^ , : ? 1[]'. Invalid character will replaced by 'X') Select SVM Threshold:	external source										
SER_example Shannon Entropy of Residues based prediction example												
CPP+SS_example (Composition Profile of Patterns + Secondary Structure) based prediction example	SVM Threshold value between ~1.00 to + 1.00. Default is 0											
DPC+SS_example (Dipeptide Composition + Secondary Structure) based prediction example	INFO: DPC+SS (Dipeptide Composition + Secondary Structure) based prediction: As sequence patterns of fixed length of 41-residues, we consider gapped dipeptides composition of both directions, where A represents an any other with white the second of 0.0 and 0.0 because the both direction of 0.1 Denian the directions, where A represents and any other with white the second of 0.0 and 0.0 because the both direction of 0.1 Denian the directions, where A represents and any other with the second of 0.0 and 0.0 because the both direction of 0.1 Denian the directions, where A represents and any other with the second of 0.0 and 0.0 because the both direction of 0.1 Denian the directions are second of 0.0 because the second of 0.0 becaus											
DPC+ASA_example (Dipeptide Composition + Accessible Surface Area) based prediction example	ammo accuments age or sector-on and an energy and accument of the sector and accument and accume											
GlycoPP2 Tools	Please provide a fasta file for GlycoPP V2.0 DPC+SS prediction tool-											

Figure 27: Example of GlycoPP V2.0 Analsis overview

Galaxy GlycoPP	V2.0	Analyze Data	Workflow Visualize	▼ Shared Data ▼ ⊦	lelp 👻 Login or Register			Using 15.0 KB
Tools	<u>1</u>	>B4EB72		Length = 171		^	History	C 🕈
search tools	0	>B4EB72	1.00	search datasets	0			
Cig Cosylucion		Potential O-Linked Glycosylate	Unnamed history					
CTD_example Conjoint Triad Descriptors Example		DSTPPDVNGFGISLKIGSVDYQN	JATEPACSAHSGPGSPPS	GKSYTLTGTGHAVIPGQ	1 shown, 6 deleted			
PAAC_example Pseudo Amino Acio Composition based prediction	ł	GlycoPP v2.0 Prediction Metho	od = DPC+SS	SVM Threshold = 0			14.98 KB	
example		Position		Residue	Score	Prediction	7: DPC+SS example	⊛ & ×
SER example Shannon Entropy of		2		т	-0.46211177	Non-alycosylated		
Residues based prediction example		7		T	-0.38791799	Non-glycosylated		
CPP+SS example (Composition Profile of Patterns + Secondary		15		T	-0.28836015	Non-glycosylated		
Structure) based prediction exampl	2	20		S	-0.31229772	Non-glycosylated		
DPC+SS example (Dipeptide		25		Т	-0.30410862	Non-glycosylated		
Composition + Secondary Structure)	30		S	-0.2775764	Non-glycosylated		
based prediction example		33		S	-0.31546779	Non-glycosylated		
DPC+ASA example (Dipeptide		37		S	-0.29199153	Non-glycosylated		
Composition + Accessible Surface		40		S	-0.22250738	Non-glycosylated		
Area) based prediction example		43		S	-0.19816666	Non-glycosylated		
GlycoPP2 Tools		49		Т	-0.20922205	Non-glycosylated		

Figure 28: Example of GlycoPP V2.0 Analsis output for DPC+SS(Dipeptide Composition + Secondary Structure)forOlinked glycosylation

DPC+ASA example (Dipeptide Composition + Accesible surface Area)based prediction example

Galaxy GlycoPP	V2.0	Analyze Data 🛛 Workflow Visualize 👻 Shared Data 👻 Help 🔻 Login or Register 🗮			Using 0 bytes
Tools search tools	± 0	DPC+ASA_example (Dipeptide Composition + Accessible Surface Area) based prediction example (Galaxy Version 0.1.1) • Options	Î	History search datasets	¢ 0 0
based prediction example	-	Preloaded example for input fasta sequence		Unnamed history	
Examples of O-Linked Glycosylation CTD_example Conjoint Triad Descriptors Example		>B4EB72 MTDPRHTVRIAVGATALGVSALGATLPACSAHSGPGSPPSAPSAPAAATVMVEGHTHTISGVVECRTSPAVRTATPSESGTQTTRVNAHDDSASVTLSLS DSTPPDVNGFGISLKIGSVDVQMPYQPVQSPTQVEATRQGKSYTLTGTGHAVIPGQTGMRELPFGVHVTCP		(empty)	u can load ata from an
PAAC_example Pseudo Amino Acid Composition based prediction example	I	Input fasta sequence (Allowed special character in fasta header line ' - = () / + * ^ , ; ? ! []'. Invalid character will replaced by 'X') Select SVM Threshold:		external source	
SER_example Shannon Entropy of Residues based prediction example		•			
CPP+SS_example (Composition Profile of Patterns + Secondary Structure) based prediction example		SVM Threshold value between -1.00 to +1.00. Default is 0			
DPC+SS_example (Dipeptide Composition + Secondary Structure) based prediction example		INFO: DPC+ASA (Dipeptide Composition + Accessible Surface Area): As sequence platterns of fixed length of 41-residues, we consider gapped dipeptides composition of both directions, where A represents an amino active burge a page 67 (Jon 11 and 12) have use not the best performance at On-1 Surface Accessibility is produced at another			
DPC+ASA_example (Dipeptide Composition + Accessible Surface Area) based prediction example		immore and, naming a gap or Q (Q-V), and a, nerve we get the vex performance at Q=1. Surface accessionly is employed as another important feature because glycosylation has tenancy to occur at extracellular regions of proteins with the side chain of serine or threonine residue in the sequon exposed to the surface.			
GlycoPP2 Tools		Please provide a fasta file for GlycoPP V2.0 DPC+ASA prediction tool:-			

Figure 29: Example of GlycoPP V2.0 Analsis overview

🕤 Galaxy GlycoPP	V2.0		Analyze Data	Workflow	Visualize 🕶	Shared Data 🕶	Help ▼	Login or Register			Using 17.3 KB
Tools	1	>B4EB72				Length = 171				History	C 🕈
search tools	8	>B4EB72		1.00		Length = 171				search datasets	8
CTD_example Conjoint Triad Descriptors Example PAAC_example Pseudo Amino Aci	đ	MTDPRHTVRIAV DSTPPDVNGFG GlycoPP v2.0 Pr	GATALGVSAL SLKIGSVDYQI	GATLPACSAF MPYQPVQSP od = DPC+A	ISGPGSPPSA TQVEATRQGI SA	PSAPAAATVMVEC (SYTLTGTGHAVIP SVM Threshold	GHTHTISG GQTGMRE d = 0	VVECRTSPAVRTATPS ELPFGVHVTCP	SESGTQTTRVNAHDDS	Unnamed history 1 shown, 7 deleted 17.27 KB	Y
composition based prediction example		Position				Residue		Score	Prediction	8: DPC+ASA example	@ 🖋 🗙
SER example Shannon Entropy of Residues based prediction example		2				Т		-0.52297174	Non-glycosylated		
CPP+SS example (Composition Profile of Patterns + Secondary		7 15				T T		-0.40102596 -0.32246258	Non-glycosylated Non-glycosylated		
Structure) based prediction example	e	20 25				S T		-0.31365379 -0.29185957	Non-glycosylated Non-glycosylated		
Composition + Secondary Structure based prediction example	e)	30 33				S S		-0.28479011	Non-glycosylated		
DPC+ASA example (Dipeptide Composition + Accessible Surface	11	37 40				S		-0.29001287	Non-glycosylated		
Area) based prediction example		43				s		-0.27122896	Non-glycosylated		
Workflows		49 56				T		-0.27632384 -0.25434264	Non-glycosylated		

Figure 30: Example of GlycoPP V2.0 Analsis output for DPC+ASA example (Dipeptide Composition + Accesible surface Area) for O-linked glycosylation

2. Glycopp V2.0 Analysis

A user can upload the fasta file using the Galaxy tool **Get Data as we shown in figure 8** this fasta files can be used for GlycoPP V2.0 Analysis. The following figure shows the file upload method.

Get Data Galaxy tool for uploading the fasta protein sequence for prediction of N- and O_linked glycosite



Figure31: GlycoPP V2.0 Analsis overview

a) N-linked Glycosylation

BPP Binary Profile Pattern based prediction

Galaxy GlycoPP V2.0	Analyze Data Workflow Visualize * Shared Data * Help * Login or Register 🗰	Using 180 bytes
Tools <u>±</u>	PPD Risses Profile Datters based acceletion (Colour Version 0.1.0)	2 ¢
search tools	• Options search dat	asets 😮
	Input Fasta File Unnamed	history
BPP Binary Profile Pattern based prediction	□ 12 Pasted Entry	
BPP+ASA (Binary Profile Composition+ Accessible Surface	Select input fasta or multifasta file 180 b	Ø
Area) based prediction	Select SYM Threshold: 1: Pasted Ent	try 🕑 🖋 🗙
BPP+SS (Binary Profile Composition	0 •	
+ secondary structure) based prediction	SVM Threshold value between -1.00 to +1.00. Default is 0	
BPP+ASA+SS (Binary Profile Composition+ Accessible Surface Area + Secondary Structure) based prediction	Execute INFO: BPP Binary Profile Pattern based prediction:	
O-Linked Glycosylation	Binary profile of pattern (BPP): In this approach, sequence patterns of fixed length of 41-residues were converted into binary form. Each residue of patterns was represented by a vector of dimension 20 (e.g. Ala by 1.0.00.0.0.0.0.0.0.0.0.0.0.0.0.0.0	
CTD Conjoint Triad Descriptors	0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0	
PAAC Pseudo Amino Acid Composition based prediction	Please provide a fasta file for GlycoPP V2.0 BPP prediction tool:-	
SER Shannon Entropy of Residues based prediction	Example input single fasta file	

Figure-32: GlycoPP V2.0 Analsis overview for BPP

Galaxy GlycoPP V	/2.0	Analyze I	Data Workflow Vis	sualize 👻 Shared Data 🔻	Help ▼ Login or Regis	ter		Using 3.5 KB
Tools	1	>Q0PAM0		Length = 171			History	C 🕈
search tools	8	>Q0PAM0	1.1.10	Length = 171			search datasets	0
Fetch Alignments/Sequences	*	MAIKIFGILIALFTITFTILSLC	2007 SLNLQTNALNFKI	NIEAKNLKAYESNTSIIKAYY	KANSWVRYADRDEFNDFIT	LNLDFNLSANRLEFFNKDM	Unnamed history	
Operate on Genomic Intervals		LFEGNVTYIGSNNVKIISQE		KALINGSIINGNTLNYDVK	NKILNIQGVNAWLQDK		1 shown, 1 deleted	
Statistics				SVW meshold = 0			3.47 KB	
Graph/Display Data		Position		Residue	Score	Prediction	2: BPP example	• / ×
Phenotype Association		29		NLQ	-0.68692359	Non-glycosylated		
GLYCOPP V2.0 ANALYSIS		33		NAL	-0.72174881	Non-glycosylated		
N-Linked Glycosylation		36		NFK	-0.12207736	Non-glycosylated		
O Linked Obcosylation		39		NIE	-0.32470975	Non-glycosylated		
O-Elliked Glycosylation		44		NLK	-0.20398314	Non-glycosylated		
EXAMPLES OF GLYCOPP V2.0		51		NTS	1.0002212	Potential Glycosylated		
ANALYSIS		62		NSW	-0.52280092	Non-glycosylated		
Examples of N-Linked		74		NDF	-0.35368478	Non-glycosylated		
Glycosylation		80		NLD	-0.8578013	Non-glycosylated		
RDD example Bipapy Profile Pattern		84		NLS	0.71732003	Potential Glycosylated		
based prediction example		88		NRL	-0.36008792	Non-glycosylated		
RDD+ ASA example (Binany Profile		94		NKD	-0.15063364	Non-glycosylated	Activate Windows	
Composition + Accessible Surface		105		NVT	0.63372566	Potential Glycosylated	Go to Settings to activat	
Area) hased prediction example	*	112		NNV	-0.70692049	Non-glycosylated	 Go to setalligs to activat 	c militadays.

Figure 33: GlycoPP V2.0 Analsis output for BPP (Binary Profile Pattern) forN-linked glycosylation

BPP+ASA (Binary Profile Composition+ Accessible Surface Area) based prediction



Figure 34: GlycoPP V2.0 Analsis

Galaxy GlycoPP V	2.0	Analyze Da	ta Workflow	Visualize 👻 Shared Data 🎙	Help 🔻 Login or Regis	ster		Using 3.5 KB
Tools	<u>t</u>	>Q0PAM0		Length = 171		A	History	C 🕈
search tools	8	>Q0PAM0		Length = 171			search datasets	8
Fetch Alignments/Sequences	^	Potential N-Linked Glycosy MAIKIFGILIALFTITFTILSLQD	lated Sites: PYSLNLQTNALNI	KNIEAKNLKAYESNTSIIKAY	'KANSWVRYADRDEFNDFIT	UNLDFNLSANRLEFFNKDM	Unnamed history	
Operate on Genomic Intervals		LFEGNVTYIGSNNVKIISQEV	eyqpkdkilhtnt	NFKALINGSIINGNTLNYDVK	NKILNIQGVNAWLQDK		1 shown, 1 deleted	
Statistics		GlycoPP v2.0 Prediction Me	thod = BPP	SVM Threshold =	0		3.47 KB	
Graph/Display Data		Position		Residue	Score	Prediction	2: BPP example	● / ×
Phenotype Association	1			NLO	-0.68692359	Non-alvcosvlated		
GLYCOPP V2.0 ANALYSIS		33		NAL	-0.72174881	Non-glycosylated		
N-Linked Glycosylation		36		NFK	-0.12207736	Non-glycosylated		
O-Linked Glycosylation		39		NIE	-0.32470975	Non-glycosylated		
O-Linked Glycosylation		44		NLK	-0.20398314	Non-glycosylated		
EXAMPLES OF GLYCOPP V2.0		51		NTS	1.0002212	Potential Glycosylated		
ANALYSIS		62		NSW	-0.52280092	Non-glycosylated		
Examples of N-Linked		74		NDF	-0.35368478	Non-glycosylated		
Glycosylation		80		NLD	-0.8578013	Non-glycosylated		
PDD evenuele Pinany Profile Pattern		84		NLS	0.71732003	Potential Glycosylated		
based prediction example		88		NRL	-0.36008792	Non-glycosylated		
PDD + ASA example (Rippo) Profile		94		NKD	-0.15063364	Non-glycosylated	Activata Windows	
Composition + Accessible Surface		105		NVT	0.63372566	Potential Glycosylated	Go to Settings to activat	
Area) based prediction example	*	112		NNV	-0.70692049	Non-glycosylated		e windows.

Figure 35: GlycoPP V2.0 Analsis output for BPP+ASA (Binary Profile Pattern+ Accessible Surface Are) forN-linked glycosylation

BPP+ASA+SS (Binary Profile Composition+ Accessible Surface Area + Secondary Structure) based prediction



Figure 36: GlycoPP V2.0 Analsis overview

Galaxy GlycoPP V2	Analyze Da	ta Workflow Visualize	Shared Data He HE	elp 👻 Login or Register			Using 3.5 KB
Tools 📩	>Q0PAM0		Length = 171		A	History	C 🕈
search tools	>Q0PAM0		Length = 171			search datasets	8
	Potential N-Linked Glycosy	ated Sites:					
Get Data	MAIKIFGILIALFTITFTILSLQD	PYSLNLQTNALNFKNIEAKN	ILKAYESNTSIIKAYYKAN	SWVRYADRDEFNDFITLNI	LDFNLSANRLEFFNKDM	Unnamed history	
Send Data	LFEGNVTYIGSNNVKIISQEVI	YQPKDKILHTNTNFKALIN	GSIINGNTLNYDVKNKILI	NIQGVNAWLQDK		2 shown	
	GlycoPP v2.0 Prediction Me	thod = BPP+ASA+SS	SVM Threshold =	0		3.5 KB	
Lift-Over							
Expression Tools	Position		Residue	Score	Prediction	2: BPP+SS example	👁 🖋 🗙
Collection Operations	29		NLQ	-0.61250663	Non-glycosylated	1: BPP+ASA+SS example	⊛ # ×
Text Manipulation	33		NAL	-0.53444309	Non-glycosylated		
Convert Formats	36		NFK	0.026839411	Potential Glycosylat		
	39		NIE	-0.29871205	Non-glycosylated		
Filter and Sort	44		NLK	-0.30010003	Non-glycosylated		
Join, Subtract and Group	51		NTS	0.58642466	Potential Glycosylat		
Fetch Alianments/Sequences	62		NSW	-0.39528659	Non-glycosylated		
reten Aighments/Sequences	74		NDF	-0.20714487	Non-glycosylated		
Operate on Genomic Intervals	80		NLD	-0.65906657	Non-glycosylated		
Statistics	84		NLS	0.55723666	Potential Glycosylat		
Graph (Display, Data	88		NRL	-0.28949737	Non-glycosylated		
Graph/Display Data	94		NKD	-0.23029647	Non-alvcosvlated	and a state of	

Figure 37 GlycoPP V2.0 Analsis output for BPP+ASA+SS (Binary Profile Composition+ Accessible Surface Area + Secondary Structure) forN-linked glycosylation

BPP+SS (Binary Profile Composition+ Secondary Structure) based prediction



Figure 38: GlycoPP V2.0 Analsis overview

Tools 1	>Q0PAM0	Length = 171		A	History	C 🕈
search tools	>QOPAMO	Length = 171			search datasets	0
operate on denomic intervais	Potential N-Linked Glycosylate MAIKIFGILIALFTITFTILSLQDPYS	d Sites: LNLQTNALNFKNIEAKNLKAYESNTSIIKAYYI	KANSWVRYADRDEFNDFITL	NLDFNLSANRLEFFNKDM	Unnamed history	
Statistics	LFEGNVTYIGSNNVKIISQEVEYQ	PKDKILHTNTNFKALINGSIINGNTLNYDVKN	IKILNIQGVNAWLQDK		2 shown	
Graph/Display Data	GlycoPP v2.0 Prediction Metho	d = BPP+SS SVM Threshold =	= 0		3.5 KB	
Phenotype Association	Position	Residue	Score	Prediction	2: PDD+SS example	
GLYCOPP V2.0 ANALYSIS					2. BPP+35 example	
N-Linked Glycosylation	29	NLQ	-0.61294694	Non-glycosylated	1: BPP+ASA+SS example	👁 🖋 🗙
O Linked Chassadation	33	NAL	-0.53504833	Non-glycosylated		
O-Linked Glycosylation	36	NFK	0.026195499	Potential Glycosylate		
EXAMPLES OF GLYCOPP V2.0	39	NIE	-0.29947346	Non-glycosylated		
ANALYSIS	44	NLK	-0.300722	Non-glycosylated		
Examples of N-Linked	51	NTS	0.58505686	Potential Glycosylate		
Glycosylation	62	NSW	-0.39570128	Non-glycosylated		
Even a f O Linked	74	NDF	-0.20760915	Non-glycosylated		
Examples of O-Linked	80	NLD	-0.65981998	Non-glycosylated		
Glycosylation	84	NLS	0.55655757	Potential Glycosylate		
GlycoPP2 Tools	88	NRL	-0.28977783	Non-glycosylated		

Figure 39 GlycoPP V2.0 Analsis output for BPP+ASA+SS (Binary Profile Composition+ Accessible Surface Area + Secondary Structure) forN-linked glycosylation

b) O-linked Glycosylation

CTD Conjoint Triad Descriptors



Figure 40: GlycoPP V2.0 Analsis overview

Salaxy Glyco	o PP V2.0		Analyze Data	Workflow V	isualize 👻 Shared	Data 👻 Help	 Login or Registe 	r 		Using 5.8 KB
Tools	<u>±</u>	>B4EB72			Length = 1	171		*	History	C 🕈
search tools	8	>B4EB72	ked Glycosylate	- Sites	Length = 1	71			search datasets	8
operate on denomic in		MTDPRHTVRIA	VGATALGVSALG	ATLPACSAHSG	PGSPPSAPSAPAAA	TVMVEGHTHTIS	GVVECRTSPAVRTAT	PSESGTQTTRVNAHDDS	Unnamed history	
Statistics		DSTPPDVNGF	SISLKIGSVDYQM	PYQPVQSPTQ	/EATRQGKSYTLTGT	GHAVIPGQTGM	RELPFGVHVTCP		1 shown, 2 deleted	
Graph/Display Data		GlycoPP v2.0 P	rediction Metho	d = CTD	SVM Thres	hold = 0			5.8 KB	
Phenotype Association		Position			Residue		Score	Prediction		0 1 4
GLYCOPP V2.0 ANALYSIS									3: CTD_example	• / ×
N-Linked Glycosylation		2			Т		-0.20935397	Non-glycosylated		
O Linked Chrocylation		7			T		-0.19029675	Non-glycosylated		
O-Linked Glycosylation		15			Т		-0.1508115	Non-glycosylated		
EXAMPLES OF GLYCOPP V2.	0	20			S		-0.40016528	Non-glycosylated		
ANALYSIS		25			т		-0.11553263	Non-glycosylated		
Examples of N-Linked		30			S		-0.13035721	Non-glycosylated		
Glycosylation		33			S		0.039000312	Potential Glycosylat		
		37			S		-0.11382714	Non-glycosylated		
Examples of O-Linked		40			S		-0.11227607	Non-glycosylated		
Glycosylation		43			S		-0.080158341	Non-glycosylated		
GlycoPP2 Tools		49			т		0.030458067	Potential Glycosylat		
-		50			т		0.102006.41	Detential Character		

Figure 41: GlycoPP V2.0 Analsis output for CTD (Conjoint Triad Distributor) for O-linked glycosylation

PAAC Pseudo Amino Acid Composition based prediction



Figure 42: GlycoPP V2.0 Analsis overview

🕤 Galaxy GlycoPP	V2.0	F	nalyze Data	Workflow	Visualize 🔻	Shared Data 🕶	Help 🔻	Login or Register			Using 8.1 KB
Tools	<u>1</u>	>B4EB72			Le	ength = 171			A	History	C 🕈
search tools	8	>B4EB72			Le	ength = 171				search datasets	0
ory cosynation		Potential O-Linke	d Glycosylate	ed Sites:	CDCCDDCAD				CECCTOTED ALAUDDO	Unnamed bistomy	
CTD_example Conjoint Triad Descriptors Example		DSTPPDVNGFGIS	LKIGSVDYQN	JATEPACSAHS IPYQPVQSPT	QVEATRQGK	SYTLTGTGHAVIPO	GQTGMRE	ELPEGVHVTCP	SESGIQITRVINAHDDS/	1 shown, 3 deleted	
PAAC_example Pseudo Amino Aci	d	GlycoPP v2.0 Pre	diction Metho	od = PAAC	S\	/M Threshold = 0				8.09 KB	
example		Position			Re	esidue		Score	Prediction	4: PAAC_example	● / ×
SER example Shannon Entropy of Residues based prediction example	2	2			т			-0.20271261	Non-glycosylated		
CPP+SS example (Composition		7			Т			-0.18153102	Non-glycosylated		
Profile of Patterns + Secondary		15			Т			-0.047441369	Non-glycosylated		
Structure) based prediction example	le	20			S			-0.15534192	Non-glycosylated		
DPC+SS example (Dipeptide		25			Т			-0.15297738	Non-glycosylated		
Composition + Secondary Structure	e)	30			S			-0.18510155	Non-glycosylated		
based prediction example		33			S			-0.18575067	Non-glycosylated		
DPC+ASA example (Dipeptide		37			S			-0.11436738	Non-glycosylated		
Composition + Accessible Surface		40			S			0.012788534	Potential Glycosyla		
Area) based prediction example		43			S			-0.082389054	Non-glycosylated		
GlycoPP2 Tools		49			Т			-0.010831194	Non-glycosylated		
					-						

Figure 43: GlycoPP V2.0 Analsis output for PAAC (Pseudo Amino Acid Composition) for O-linked glycosylation

SER Shannon Entropy of Residues based prediction

Galaxy GlycoPP V2.0	Analyze Data Workflow Visualize * Shared Data * Help * Login or Register	Using 1.9 KB
Tools 🛓	FIR Character Education (Calco Vicine 010)	C 🕈
search tools	Sex Shannon entropy or Residues based prediction (Galaxy Version 0.1.0) • Options search datasets	0
н силкса отусозуластотт	FASTA file	
BPP Binary Profile Pattern based	Unitaria de la constance de la constan Constance de la constance de	
prediction	C 42 C 1: Pasted Entry	
BPP+ASA (Binary Profile	Select fasta file 1.91 KB	
Area) based prediction	Select SVM Threshold: 1: Pasted Entry	⊛ 🖋 X
BPP+SS (Binary Profile Composition	0 -	
+ Secondary Structure) based prediction	SVM Threshold value between -1.00 to +1.00. Default is 0	
BPP+ASA+SS (Binary Profile	✔ Execute	
Area + Secondary Structure) based		
prediction	INFO: SER (Shannon Entropy of Residues):	
O-Linked Glycosylation	To understand the structural orchestration of sequences i.e., propensity towards order and disorder. In this regard, the Shannon entropy (SE) score was calculated for each consensus sequence. As it was evidenced that entropy possesses an idea of the disorder. Entropy was	
CTD Conjoint Triad Descriptors	directly proportional to the rate of disorder i.e., if the disorder increases, it signifies higher entropy.	
PAAC Pseudo Amino Acid Composition based prediction	Please provide a fasta file for GlycoPP V2.0 SER prediction tool-	
SER Shannon Entropy of Residues	Example input single fasta file	
based prediction	>846872	
CPP+SS (Composition Profile of	MTPPRHTVRIAVGATALGVSALGATLPACSAHSGPGSPPSAPSAPAARTVMVEGHTHTISGVVECRTSPAVRTATPSESGTQTTRVNAHDDSASVTLSLSDSTPPDVNGFGISLKIGSVDYQNPYQPVQSP	

Figure 44: GlycoPP V2.0 Analsis overview

Galaxy GlycoPP	V2.0	Anal	yze Data W	orkflow Visua	alize 🔻 Shared Data	▪ Help ▪	Login or Registe	r 💶		Using 10.4 KB
Tools	<u>*</u>	>B4EB72			Length = 171			Î	History	C 0
search tools	0	>B4EB72	lucosulated Si	tor	Length = 171				search datasets	8
Get Data		MTDPRHTVRIAVGAT	ALGVSALGATI	PACSAHSGPG	SPPSAPSAPAAATVMV	EGHTHTIS	GVVECRTSPAVRTAT	PSESGTQTTRVNAHDDS	Unnamed history	
Send Data		DSTPPDVNGFGISLKI	GSVDYQMPY	QPVQSPTQVEA	TRQGKSYTLTGTGHAV	/IPGQTGMF	ELPFGVHVTCP	_	1 shown, 4 deleted	
Lift-Over		Giycopp vz.0 predicti	ion Method =	5EK	SVIVI Inreshold	= 0			10.38 KB	
Expression Tools		Position			Residue		Score	Prediction	5: SER example	@ 🖋 🗙
Collection Operations		2			т		-0.36043891	Non-glycosylated		
Text Manipulation		7			Т		-0.5362917	Non-glycosylated		
Convert Formats		15			Т		-0.29736866	Non-glycosylated		
Filter and Sort		20 25			S T		-0.1583557 -0.18603931	Non-glycosylated Non-glycosylated		
Join, Subtract and Group	- 1	30			S		-0.13741532	Non-glycosylated		
Fetch Alianments/Sequences		33			S		-0.16468103	Non-glycosylated		
reten Anglintents/Sequences		37			S		-0.10327306	Non-glycosylated		
Operate on Genomic Intervals	5	40			S		0.1693945	Potential Glycosylate		
Statistics		43			S		0.044048193	Potential Glycosylate		
Graph/Display Data		49			Т		0.066383281	Potential Glycosylate		

Figure 45: GlycoPP V2.0 Analsis output for SER (Shannon Entropy of Residues) for O-linked glycosylation

CPP+SS (Composition Profile of Patterns + Secondary Structure) based prediction

Galaxy GlycoPP V2.0	Analyze Data 🛛 Workflow Visualize 🔻 Shared Data 🍷 Help 🔻 Login or Register 🗮	Using 1.9 KB
Tools 🛓	CPPLASS (Composition Details of Deterror + Secondary Structure) haved prediction (Calava Varrian 0.1.0)	C 🕈
search tools	CT + SS (composition Frome of Facterins + Secondary Structure) based prediction (calaxy version on to)	0
prediction	Input Fasta File Unnamed history	
O-Linked Glycosylation	🖸 🖗 🗅 1: Pasted Entry 👻 🕞 1 shown, 1 deleted	
CTD Conjoint Triad Descriptors	Select innut facta or multifacta file 1.91 KB	
PAAC Pseudo Amino Acid Composition based prediction	Select SVM Threshold: 1: Pasted Entry	@ # ¥
SER Shannon Entropy of Residues based prediction		
CPP+SS (Composition Profile of Patterns + Secondary Structure) based prediction	SVM Threshold value between -1.00 to +1.00. Default is 0	
DPC+SS (Dipeptide Composition + Secondary Structure) based prediction	INFO: CPP+SS (Composition Profile of Patterns + Secondary Structure) based prediction :	
DPC+ASA (Dipeptide Composition + Accessible Surface Area) based prediction	Composition profile of patterns is the percentage frequencies of each amino acid in a fixed length sequence patterns. Previous studies on eukaryotic glycoproteins suggested that the probability of finding glycosite was higher at positions where there was a secondary structure change. So, we applied the secondary structure along with DPC combination to get the higher result.	
EXAMPLES OF GLYCOPP V2.0 ANALYSIS	Please provide a fasta file for GlycoPP V2.0 SER prediction tool-	
Examples of N-Linked	Example input single fasta file	
Glycosylation	>846872	
Examples of O-Linked	NTDPRHTVRIAVGATALGVSALGATLPACSAHSGPGSPPSAPSAPAAATVWVEGHTHTISGWECKTSPAVRTATPSESGTQTTRVNAHDOSASVTLSLGSTPPDVWGFGISLKIGSVDYQHPYQPVQSP	

Figure 46: GlycoPP V2.0 Analsis overview

S Galaxy GlycoPP	V2.0	Analyze Data Workf	ow Visualiz	e 🔹 Shared Data 👻	Help Login or Register			Using 12.7 KB
Tools	<u>1</u>	>B4EB72		Length = 171		<u>^</u>	History	C 🕈
search tools	8	>B4EB72 Potential O Linked Chronylated Siter		Length = 171			search datasets	8
CTD_example Conjoint Triad	^	MTDPRHTVRIAVGATALGVSALGATLPAC DSTPPDVNGFGISLKIGSVDYQMPYQPV	SAHSGPGSPF QSPTQVEATR	PSAPSAPAAATVMVEGH QGKSYTLTGTGHAVIPG	HTHTISGVVECRTSPAVRTATPS	SESGTQTTRVNAHDDS	Unnamed history 1 shown, 5 deleted	
PAAC_example Pseudo Amino Aci Composition based prediction	d	GlycoPP v2.0 Prediction Method = CPI	+SS	SVM Threshold =			12.68 KB	
example		Position		Residue	Score	Prediction	6: CPP+SS example	👁 🖋 🗙
SER example Shannon Entropy of Residues based prediction example	2	2		T	-0.29620787	Non-glycosylated		
CPP+SS example (Composition		7		Т	-0.52232343	Non-glycosylated		
Profile of Patterns + Secondary		15		Т	-0.39875222	Non-glycosylated		
Structure) based prediction examp	le	20		S	-0.19466643	Non-glycosylated		
DPC+SS example (Dipeptide		25		Т	-0.20355421	Non-glycosylated		
Composition + Secondary Structur	e)	30		S	-0.20582699	Non-glycosylated		
based prediction example		33		S	-0.20324	Non-glycosylated		
DPC+ASA example (Dipeptide		37		S	-0.10289745	Non-glycosylated		
Composition + Accessible Surface		40		S	0.068873432	Potential Glycosyla		
Area) based prediction example		43		S	-0.051108665	Non-glycosylated		
GlycoPP2 Tools		49		Т	0.0087108772	Potential Glycosyla		
				-				

Figure 47: GlycoPP V2.0 Analsis output for CPP+SS (Composition Profile of Patterns + Secondary Structure) for O-linked glycosylation

DPC+SS (Dipeptide Composition + Secondary Structure)based prediction

Galaxy GlycoPP	V2.0	Analyze Data Workflow Visualize * Shared Data * Help * Login or Register		Using 1.9 KB
Tools	±.	DDC+CC /Dispatide Comparition + Secondary Structure) based prediction (Calapa Version 0.1.0)	History	C 🕈
search tools	8	Options Crast (dipeptide composition + secondary structure) based prediction (calaxy version c.n.o) Crast (dipeptide composition + secondary structure) based prediction (calaxy version c.n.o)	search datasets	8
prediction		Input Fasta File	Unnamed history	
O-Linked Glycosylation		The fin Charles Intervention	1 shown, 1 deleted	
CTD Conjoint Triad Descriptors			1.91 KB	
PAAC Pseudo Amino Acid Composition based prediction		Select input fasta or multifasta file Select SVM Threshold:	1: Pasted Entry	@ # ¥
SER Shannon Entropy of Residues based prediction		•	,	0 9 4
CPP+SS (Composition Profile of Patterns + Secondary Structure) based prediction		SVM Threshold value between -1.00 to +1.00. Default is 0		
DPC+SS (Dipeptide Composition + Secondary Structure) based prediction	on	1 INFO: DPC+SS (Dipeptide Composition + Secondary Structure) based prediction:		
DPC+ASA (Dipeptide Composition Accessible Surface Area) based prediction	•	As sequence patterns of fixed length of 41-residues, we consider gapped dipeptides composition of both directions, where A represents an amino acid, having a gap of Q (Q=0,1 and 2), here we get the best performance at Q=1. Previous studies on eukaryotic glycoproteins suggested that the probability of finding glycosite was higher at positions where there was a secondary structure change. So, we applied the secondary structure along with DPC combination to get the higher result.		
EXAMPLES OF GLYCOPP V2.0 ANALYSIS		Please provide a fasta file for GlycoPP V2.0 DPC+SS prediction tool-		
Examples of N-Linked		Example input single fasta file		
Glycosylation		>B4E872		

Figure 48: GlycoPP V2.0 Analsis overview

Galaxy GlycoPP	V2.0	Analyze Data Workflow Visu	ualize 👻 Shared Data 🍷 H	lelp 👻 Login or Register			Using 15.0 KB
Tools	£	>B4EB72	Length = 171		*	History	C 🗘
search tools	8	>B4EB72	Length = 171			search datasets	8
CTD_example Conjoint Triad Descriptors Example	•	MTDPRHTVRIAVGATALGVSALGATLPACSAHSGPC DSTPPDVNGFGISLKIGSVDYQMPYQPVQSPTQVE	SSPPSAPSAPAAATVMVEGHT ATRQGKSYTLTGTGHAVIPGQ	HTISGVVECRTSPAVRTATP TGMRELPFGVHVTCP	SESGTQTTRVNAHDDS	Unnamed history 1 shown, 6 deleted	
PAAC_example Pseudo Amino Acid Composition based prediction	ł	GlycoPP v2.0 Prediction Method = DPC+SS	SVM Threshold = 0		Prodiction	14.98 KB	
example SER example Shannon Entropy of Residues based prediction example		2	T	-0.46211177	Non-glycosylated	7: DPC+SS example	● <i>≱</i> ×
CPP+SS example (Composition Profile of Patterns + Secondary Structure) based prediction example	e	7 15 20	T T S	-0.38791799 -0.28836015 -0.31229772	Non-glycosylated Non-glycosylated Non-glycosylated		
DPC+SS example (Dipeptide Composition + Secondary Structure based prediction example	e)	25 30 33	T S S	-0.30410862 -0.2775764 -0.31546779	Non-glycosylated Non-glycosylated Non-glycosylated		
DPC+ASA example (Dipeptide Composition + Accessible Surface Area) based prediction example		37 40 43	s s	-0.29199153 -0.22250738 -0.19816666	Non-glycosylated Non-glycosylated		
GlycoPP2 Tools		49	T	-0.20922205	Non-glycosylated		

Figure 49: GlycoPP V2.0 Analsis output for DPC+SS(Dipeptide Composition + Secondary Structure) for O-linked glycosylation

DPC+ASA (Dipeptide Composition + Accesible surface Area)based prediction



Figure 50: GlycoPP V2.0 Analsis overview

Salaxy GlycoPP	V2.0	Analyze Data	Workflow Visu	alize 🔹 Shared Data 🍷 I	Help 👻 Login or Register			Using 17.3 KB
Tools	1	>B4EB72		Length = 171		ŕ	History	C 🕈
search tools	8	>B4EB72		Length = 171			search datasets	0
CTD_example Conjoint Triad Descriptors Example PAAC_example Pseudo Amino Acid Composition based prediction example	1	Potential U-Linked Giycosylated Sites: MTDPRHTVRIAVGATALGVSALGATLPACSAHSGPGSPPSAPSAPAAATVMVEGHTHTISGVVECRTSPAVRTATPSESGTQTTRVNAHDDS, DSTPPDVNGFGISLKIGSVDYQMPYQPVQSPTQVEATRQGKSYTLTGTGHAVIPGQTGMRELPFGVHVTCP GlycoPP v2.0 Prediction Method = DPC+ASA SVM Threshold = 0				Unnamed history 1 shown, 7 deleted 17.27 KB	V	
	-	Position		Residue	Score	Prediction	8: DPC+ASA example	● # ×
SER example Shannon Entropy of Residues based prediction example		2		т	-0.52297174	Non-glycosylated		
CPP+SS example (Composition Profile of Patterns + Secondary Structure) based prediction example DPC+SS example (Dipeptide Composition + Secondary Structure) based prediction example		7 15		T T	-0.40102596 -0.32246258	Non-glycosylated Non-glycosylated		
	e	20 25		S T	-0.31365379 -0.29185957	Non-glycosylated Non-glycosylated		
	<u>=</u>)	30		s	-0.28479011	Non-glycosylated		
DPC+ASA example (Dipeptide Composition + Accessible Surface Area) based prediction example		37		S	-0.29001287	Non-glycosylated		
		43		S	-0.27122896	Non-glycosylated		
Workflowe		56		T	-0.25434264	Non-glycosylated	A stimute VA/in al source	

Figure 51: GlycoPP V2.0 Analsis output for DPC+ASA (Dipeptide Composition + Accesible surface Area) for O-linked glycosylation

III. Detail Panel

This panel displays the interface of all the tools along with Input Parameters required to run a tool. It also provides help and examples to run a tool. This panel also displays the Output of a tool after its execution when user clicks on the eye so icon show in **History Panel**.

IV. History Panel

When data is uploaded from your computer or analysis is done on existing data using Galaxy, each output from those steps generates a dataset. These datasets (and the output datasets from later analysis on them) are stored by Galaxy in **Histories**.

Users that have registered an account and logged in can have many histories and the history panel allows switching between them and creating new ones.

