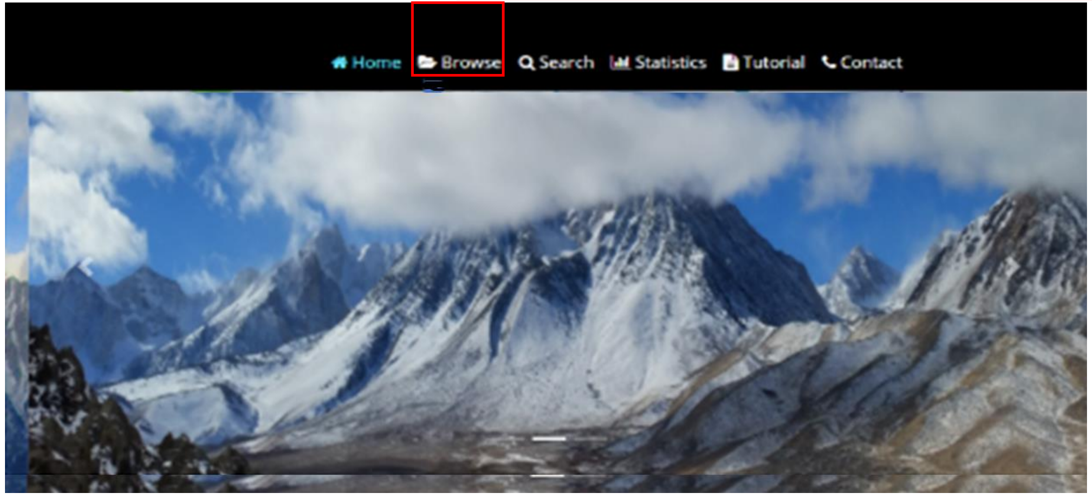
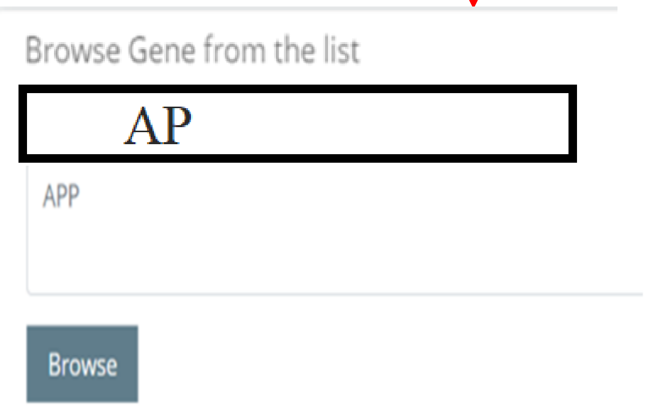
HighAltitudeOmicsDB have three browsing options



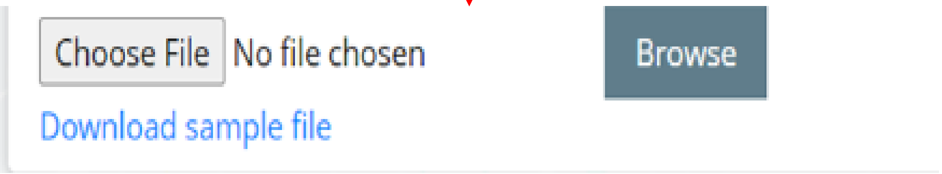
1. Browse from the database list



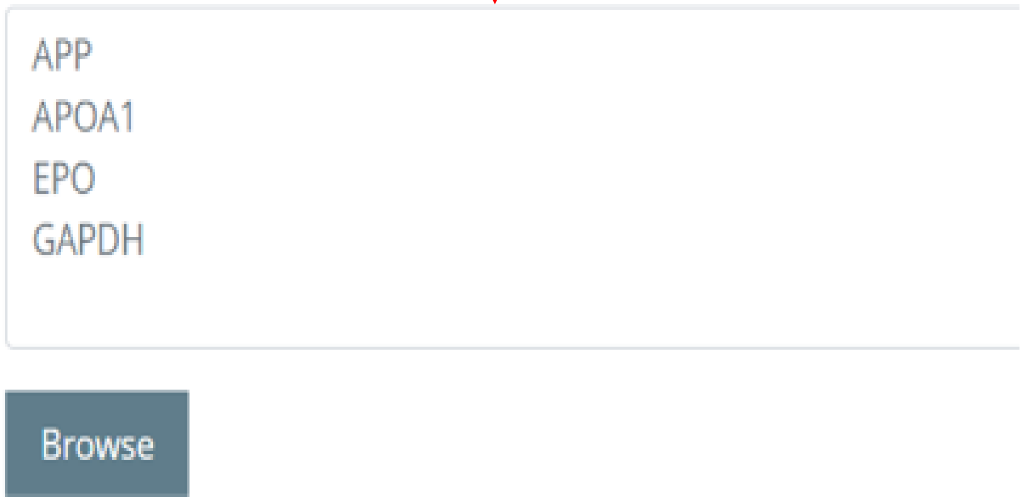
AP

APP

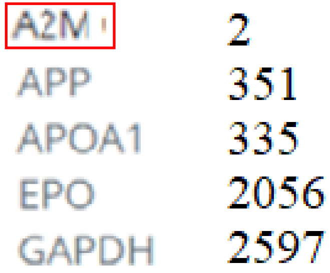
2. Uploading users list in a given format of sample file.



3. Copy and paste the list in the box as per example format shown here.



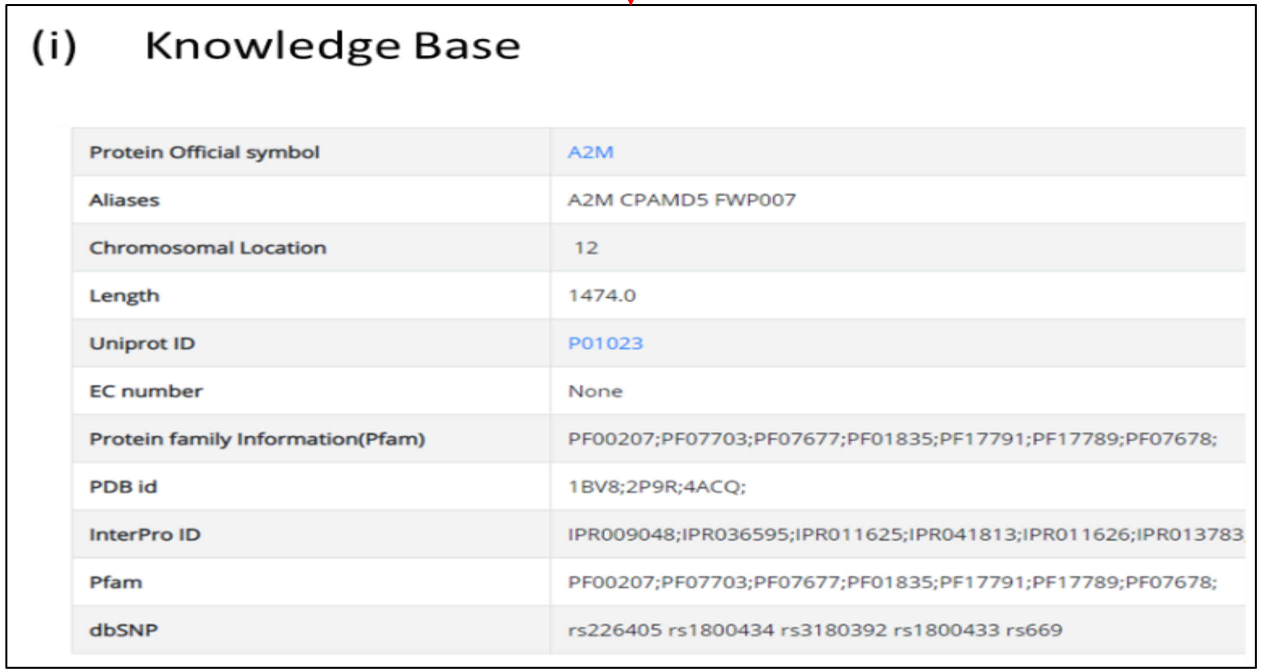
APP
APOA1
EPO
GAPDH



A2M	2
APP	351
APOA1	335
EPO	2056
GAPDH	2597

4. Browsing the list using any three options would lead to table as shown in example here.

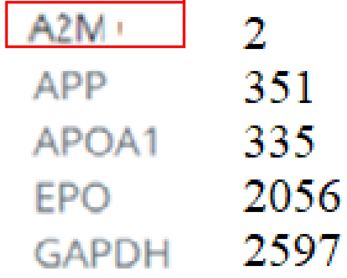
5. Selecting any protein from list as shown in example would direct it to its information page that is divided into six sections.



(i) Knowledge Base

Protein Official symbol	A2M
Aliases	A2M CPAMD5 FWP007
Chromosomal Location	12
Length	1474.0
Uniprot ID	P01023
EC number	None
Protein family Information(Pfam)	PF00207;PF07703;PF07677;PF01835;PF17791;PF17789;PF07678;
PDB id	1BV8;2P9R;4ACQ;
InterPro ID	IPR009048;IPR036595;IPR011625;IPR041813;IPR011626;IPR013783
Pfam	PF00207;PF07703;PF07677;PF01835;PF17791;PF17789;PF07678;
dbSNP	rs226405 rs1800434 rs3180392 rs1800433 rs669

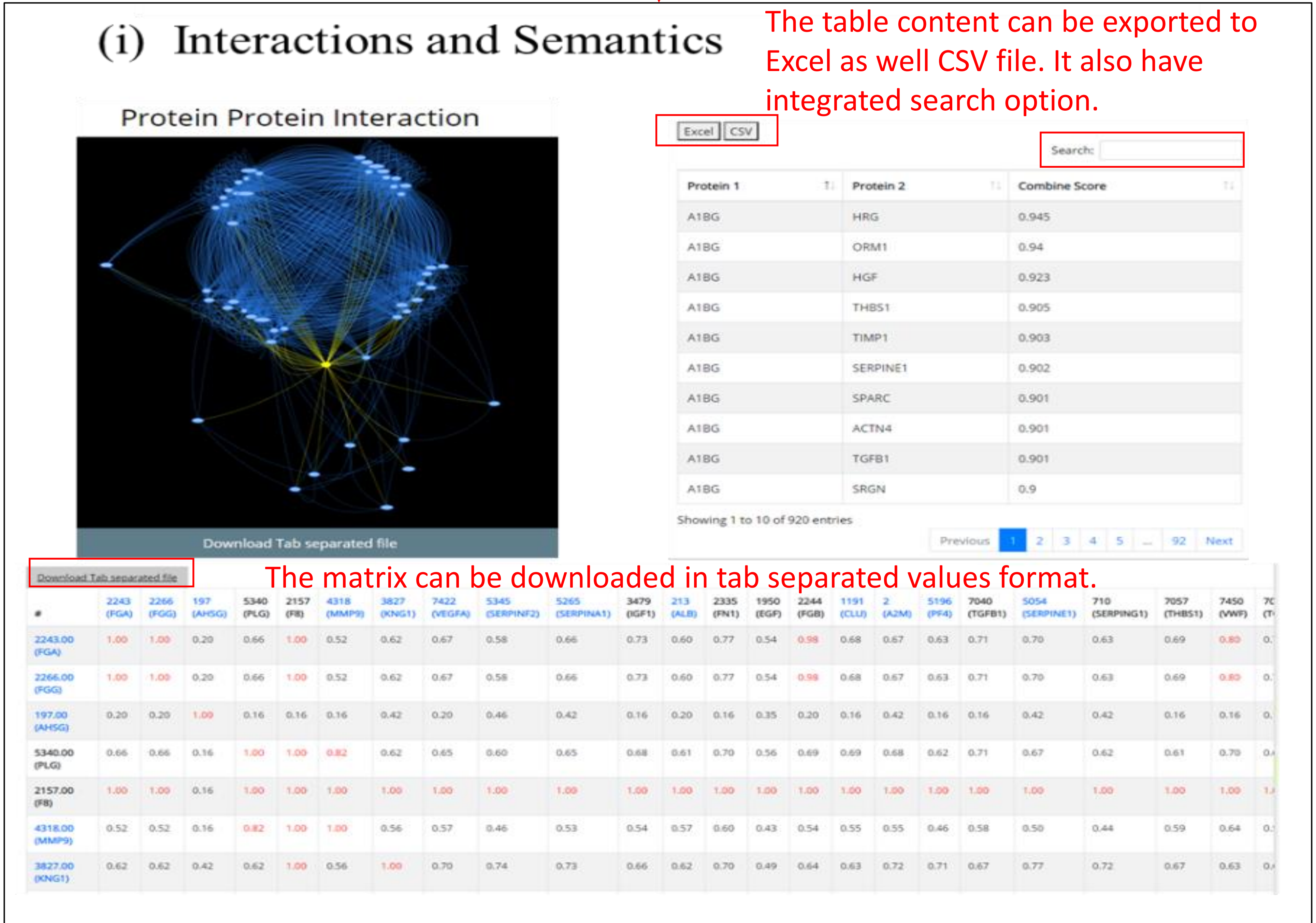
6. First Section of the information page displays knowledge base information of the proteins such as any aliases, chromosomal location, Sequence length, Uniprot ID, Enzyme Commission Number, Protein family Information (Pfam) Id, Protein Databank (PDB) Id, The Integrative Protein Signature Database (InterPro) Id, Single Nucleotide Polymorphism Database (dbSNP) Id.



A2M	2
APP	351
APOA1	335
EPO	2056
GAPDH	2597

7.Second Section displays Protein-protein interaction network of 50 closely associated interactors known from variety of sources such as experimental data, computational prediction methods, public text collections etc.

8. Additionally this section displays calculated gene semantic similarity matrix between 51 interactors to asses similarity through common molecular function.
Its also highlights scores of genes that are already present in HighAltitudeOmicsDB database



(i) Interactions and Semantics

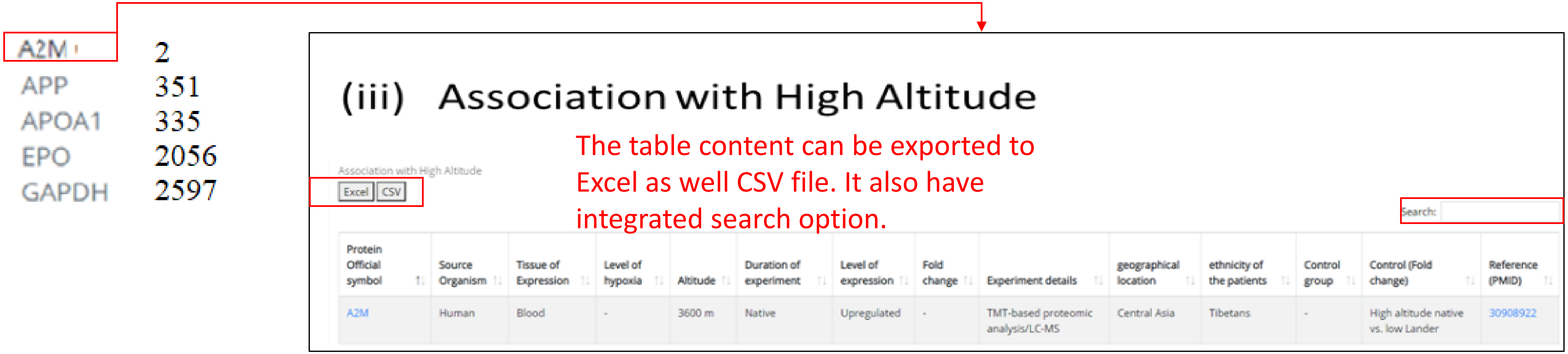
The table content can be exported to Excel as well CSV file. It also have integrated search option.

Protein 1	Protein 2	Combine Score
A1BG	HRG	0.945
A1BG	ORM1	0.94
A1BG	HGF	0.923
A1BG	THBS1	0.905
A1BG	TIMP1	0.903
A1BG	SERPINE1	0.902
A1BG	SPARC	0.901
A1BG	ACTN4	0.901
A1BG	TGFB1	0.901
A1BG	SRGN	0.9

Showing 1 to 10 of 920 entries

The matrix can be downloaded in tab separated values format.

#	2243 (FGA)	2266 (FGG)	197 (AHSG)	5340 (PLG)	2157 (F8)	4318 (MMP9)	3827 (KNG1)	7422 (VEGFA)	5345 (SERPINF2)	5265 (SERPINA1)	3479 (IGF1)	213 (ALB)	2335 (FN1)	1950 (EGF)	2244 (FGB)	1191 (CLU)	2 (A2M)	5196 (PF4)	7040 (TGFB1)	5054 (SERPINE1)	710 (SERPING1)	7057 (THBS1)	7450 (VWF)	70 (T
2243.00 (FGA)	1.00	1.00	0.20	0.66	1.00	0.52	0.62	0.67	0.58	0.66	0.73	0.60	0.77	0.54	0.98	0.68	0.67	0.63	0.71	0.70	0.63	0.69	0.80	0.
2266.00 (FGG)	1.00	1.00	0.20	0.66	1.00	0.52	0.62	0.67	0.58	0.66	0.73	0.60	0.77	0.54	0.98	0.68	0.67	0.63	0.71	0.70	0.63	0.69	0.80	0.
197.00 (AHSG)	0.20	0.20	1.00	0.16	0.16	0.16	0.42	0.20	0.46	0.42	0.16	0.20	0.16	0.35	0.20	0.16	0.42	0.16	0.16	0.42	0.42	0.16	0.16	0.
5340.00 (PLG)	0.66	0.66	0.16	1.00	1.00	0.82	0.62	0.65	0.60	0.65	0.68	0.61	0.70	0.56	0.69	0.69	0.68	0.62	0.71	0.67	0.62	0.61	0.70	0.
2157.00 (F8)	1.00	1.00	0.16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.
4318.00 (MMP9)	0.52	0.52	0.16	0.82	1.00	1.00	0.56	0.57	0.46	0.53	0.54	0.57	0.60	0.43	0.54	0.55	0.55	0.46	0.58	0.50	0.44	0.59	0.64	0.
3827.00 (KNG1)	0.62	0.62	0.42	0.62	1.00	0.56	1.00	0.70	0.74	0.73	0.66	0.62	0.70	0.49	0.64	0.63	0.72	0.71	0.67	0.77	0.72	0.67	0.63	0.



A2M	2
APP	351
APOA1	335
EPO	2056
GAPDH	2597

(iii) Association with High Altitude

The table content can be exported to Excel as well CSV file. It also have integrated search option.

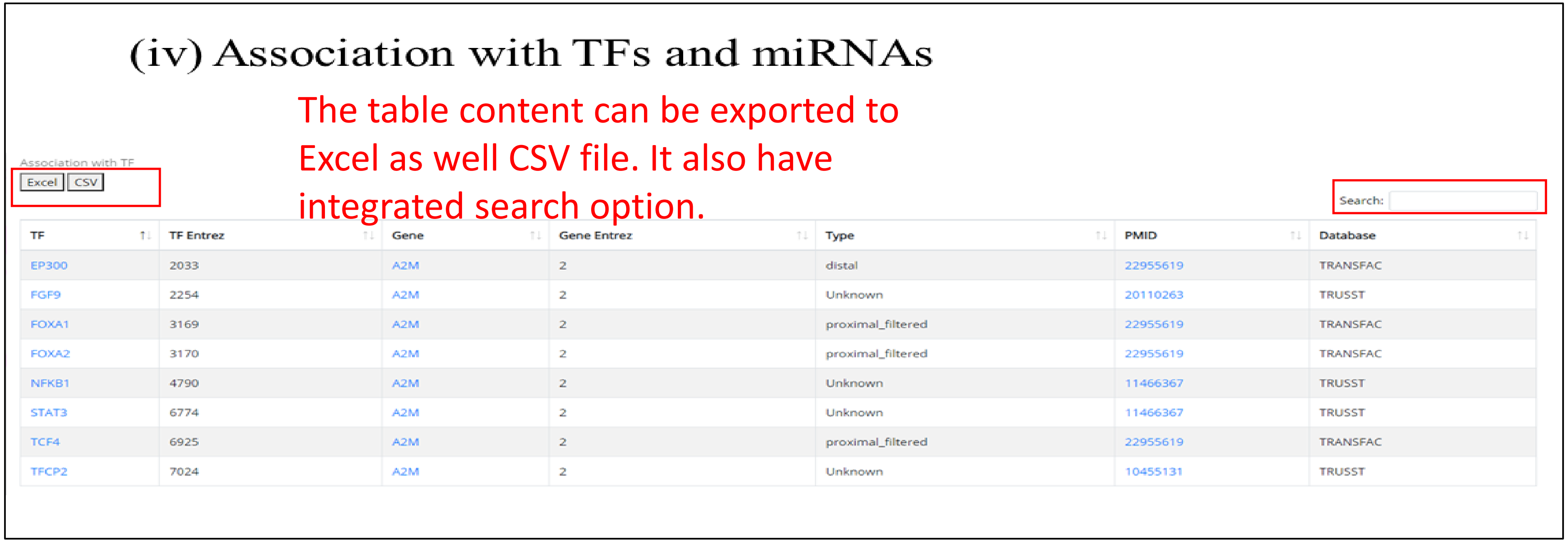
Protein Official symbol	Source Organism	Tissue of Expression	Level of hypoxia	Altitude	Duration of experiment	Level of expression	Fold change	Experiment details	geographical location	ethnicity of the patients	Control group	Control (Fold change)	Reference (PMID)
A2M	Human	Blood	-	3600 m	Native	Upregulated	-	TMT-based proteomic analysis/LC-MS	Central Asia	Tibetans	-	High altitude native vs. low Lander	30908922

9. Third section displays protein association with HA stress. The details presented in table are human protein symbol, source organism, tissue of expression, Level of Hypoxia, Altitude, Duration of the experiment, Level of expression, Fold Change, Experiment Details, Geographical location, Ethnicity etc. The short describtion of these feature are as follows.

1. The level of hypoxia is for those studies that uses artificially create hypobaric hypoxia in lab environment with similar oxygen partial pressure that existed at that altitude.
2. Source organism could any mammalian and non-mammalian system from human, lab animals, wild type, Avies etc.
3. Tissue of expression describe which specific tissue was selected for experimental study
4. The Altitude of studies incorporated in database ranges from 2500 m (medium Altitude) to ~7000 m (Extremely high altitude).
5. Duration exposure ranges from few hours to months and years. Also there is separate category for natives
6. . Level of expression could be describe in three categories i.e upregulated, downregulate and no change.
7. The experiment details can be categorize in three genomics, proteomics and transcriptomics.
8. Population Ethnicity for human study
9. Control group description and its fold change.

A2M	2
APP	351
APOA1	335
EPO	2056
GAPDH	2597

10.Fourth section list of TF and miRNAs that are known to regulate the display protein. The TFs were source from primary repositories i.e. TFcheckpoint, DBD, TcoF-DBV2, TRANSFAC databases. Whereas associated list of miRNAs have been source from MirTarBase and miRecords databases respectively.



(iv) Association with TFs and miRNAs

The table content can be exported to Excel as well CSV file. It also have integrated search option.

TF	TF Entrez	Gene	Gene Entrez	Type	PMID	Database
EP300	2033	A2M	2	distal	22955619	TRANSFAC
FGF9	2254	A2M	2	Unknown	20110263	TRUSST
FOXA1	3169	A2M	2	proximal_filtered	22955619	TRANSFAC
FOXA2	3170	A2M	2	proximal_filtered	22955619	TRANSFAC
NFKB1	4790	A2M	2	Unknown	11466367	TRUSST
STAT3	6774	A2M	2	Unknown	11466367	TRUSST
TCF4	6925	A2M	2	proximal_filtered	22955619	TRANSFAC
TFCP2	7024	A2M	2	Unknown	10455131	TRUSST

11.Fifth section list of Gene ontology and KEGG pathway association source from DAVID and KEGG mapper tools.

(v) Gene Ontology and KEGG Pathway annotations

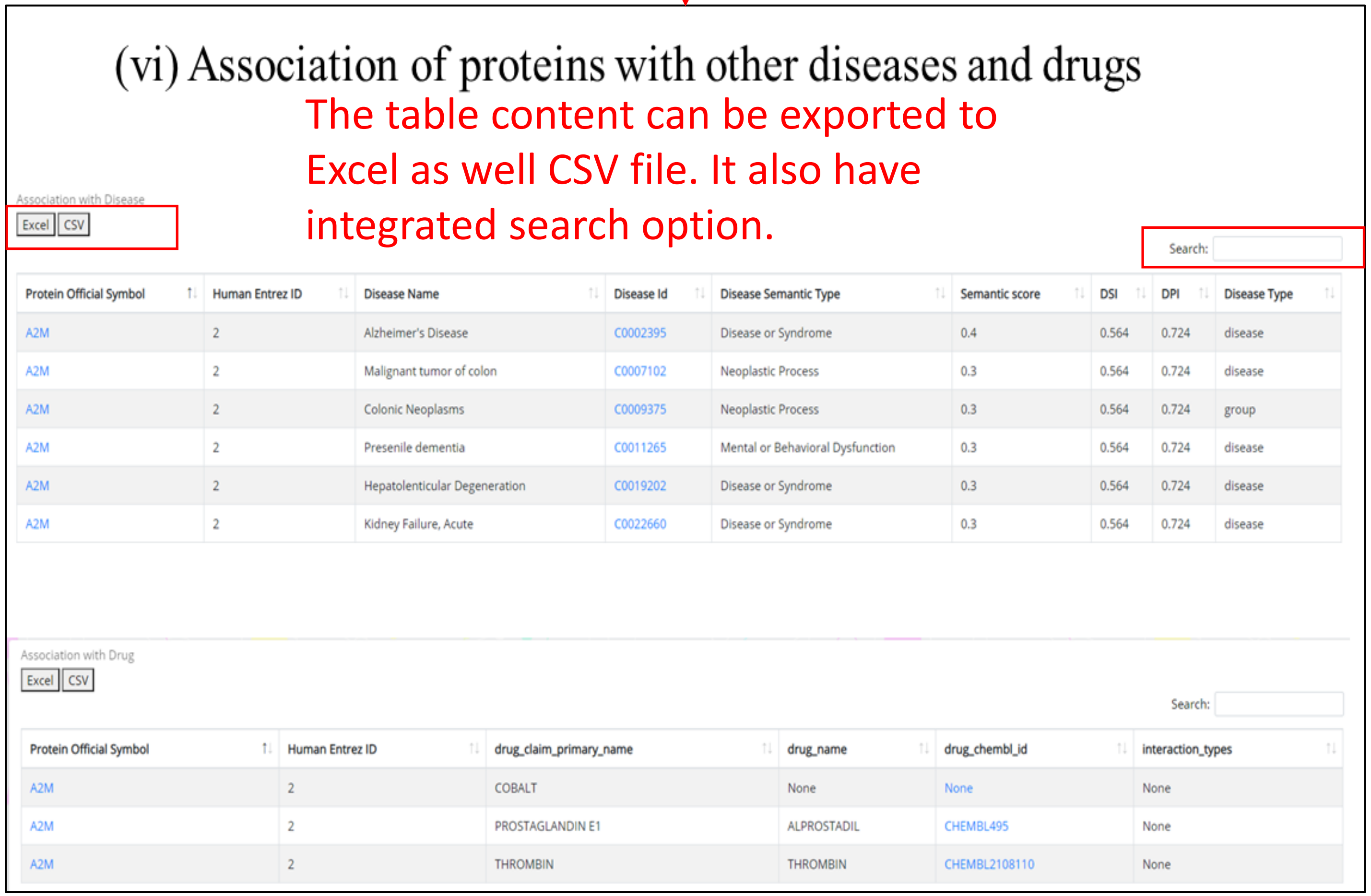
The table content can be exported to Excel as well CSV file. It also have integrated search option.

ID	GO ID	GO Term	GO Type
2	GO:0001869	negative regulation of complement activation	GOTERM_BP_DIRECT
2	GO:0002020	protease binding	GOTERM_MF_DIRECT
2	GO:0002576	platelet degranulation	GOTERM_BP_DIRECT
2	GO:0004866	endopeptidase inhibitor activity	GOTERM_MF_DIRECT
2	GO:0004867	serine-type endopeptidase inhibitor activity	GOTERM_MF_DIRECT
2	GO:0005096	GTPase activator activity	GOTERM_MF_DIRECT

Human Entrez ID	KEGG ID	KEGG Term
2	hsa04610	Complement and coagulation cascades

A2M	2
APP	351
APOA1	335
EPO	2056
GAPDH	2597

12. Sixth section displays the list of associated diseases and drugs. The list are sourced from DisGeNET and DGIdb 3.0 databases. The drug –protein association can help to show therapeutic importance of this protein. If this protein is drugable target with existing drugs or not. Similarly disease association of any target shows that abnormality in expression could be responsible for maladies in different disease that is tissue specific. Help in screening for important candidate for biomarker comparing with those that don't express in any maladies.



(vi) Association of proteins with other diseases and drugs

The table content can be exported to Excel as well CSV file. It also have integrated search option.

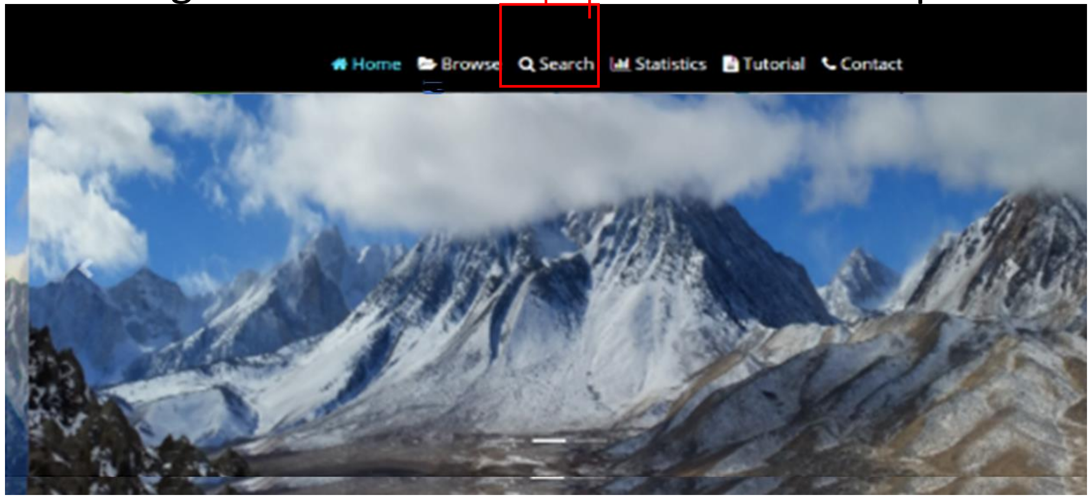
Association with Disease

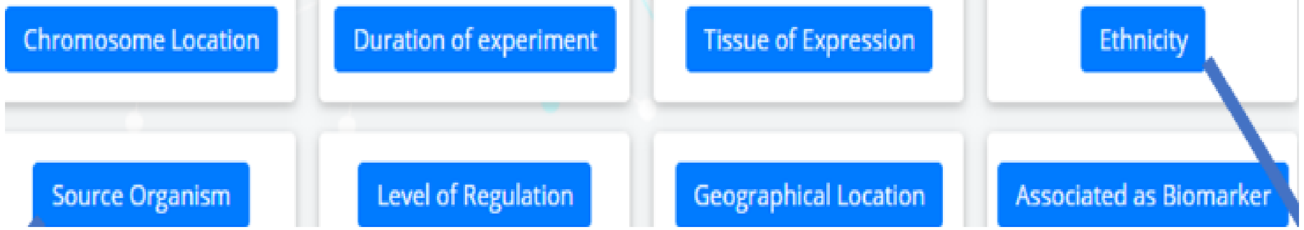
Protein Official Symbol	Human Entrez ID	Disease Name	Disease Id	Disease Semantic Type	Semantic score	DSI	DPI	Disease Type
A2M	2	Alzheimer's Disease	C0002395	Disease or Syndrome	0.4	0.564	0.724	disease
A2M	2	Malignant tumor of colon	C0007102	Neoplastic Process	0.3	0.564	0.724	disease
A2M	2	Colonic Neoplasms	C0009375	Neoplastic Process	0.3	0.564	0.724	group
A2M	2	Presenile dementia	C0011265	Mental or Behavioral Dysfunction	0.3	0.564	0.724	disease
A2M	2	Hepatolenticular Degeneration	C0019202	Disease or Syndrome	0.3	0.564	0.724	disease
A2M	2	Kidney Failure, Acute	C0022660	Disease or Syndrome	0.3	0.564	0.724	disease

Association with Drug

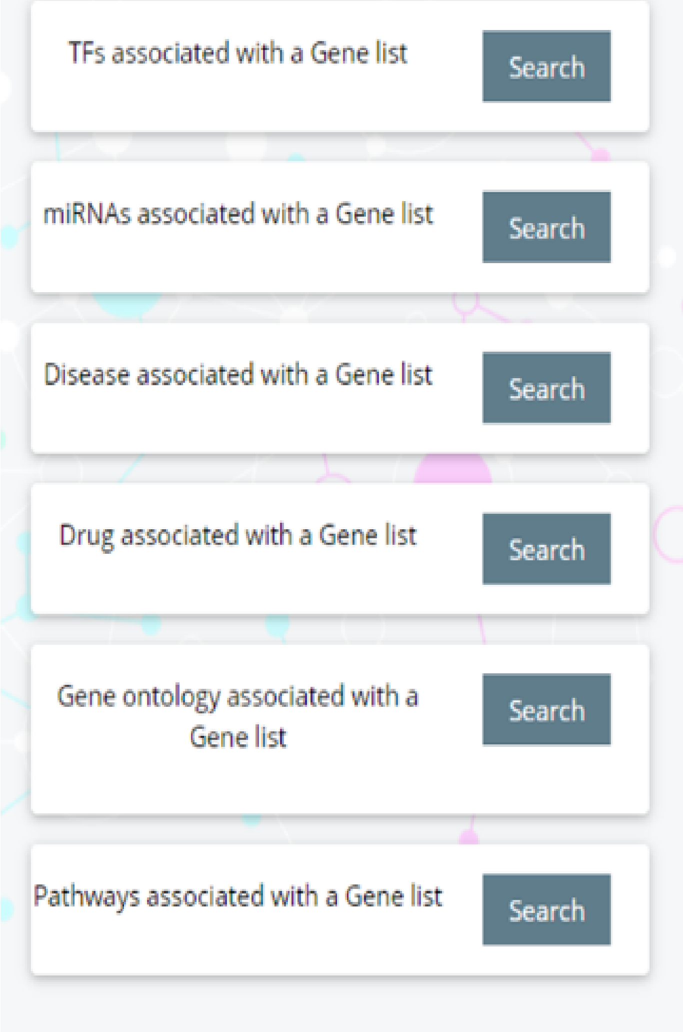
Protein Official Symbol	Human Entrez ID	drug_claim_primary_name	drug_name	drug_chembl_id	interaction_types
A2M	2	COBALT	None	None	None
A2M	2	PROSTAGLANDIN E1	ALPROSTADIL	CHEMBL495	None
A2M	2	THROMBIN	THROMBIN	CHEMBL2108110	None

HighAltitudeOmicsDB have two search options

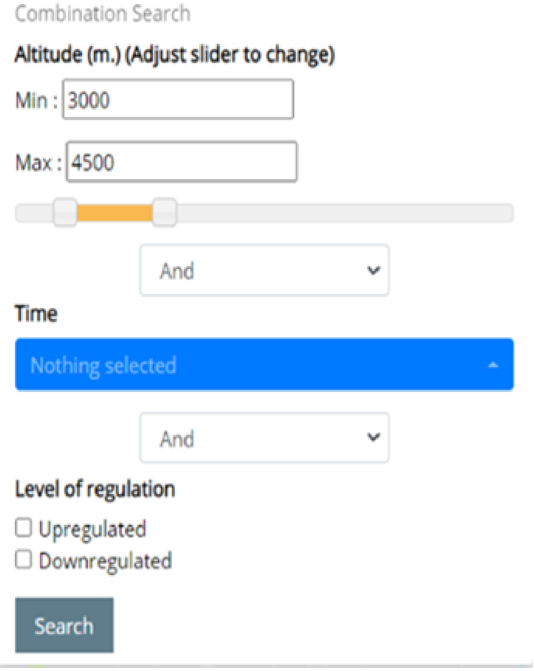




13. First search option is based on single feature search. That again divided into direct high altitude associated feature such as chromosome location, Duration of experiment, tissue of expression, ethnicity (human studies), source organism, level of regulation , geographical location, biomarker association. Second group is indirect associated features such as associated list of TFs, miRNAs, Diseases, Drugs, Gene Ontology and KEGG pathways.

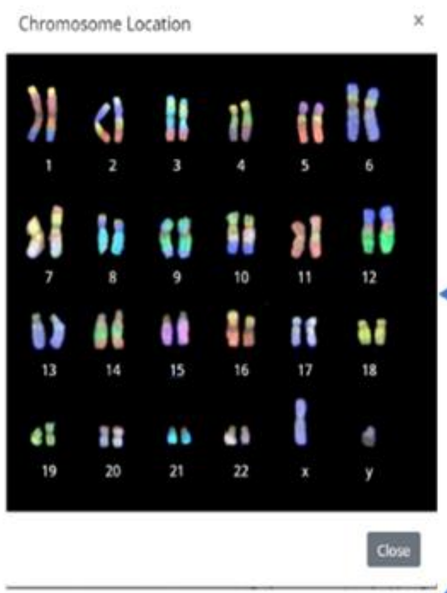


14. Second search option is combination search where features such altitude, time, level of regulation could be used in the different combination to filter out specific dataset of proteins.



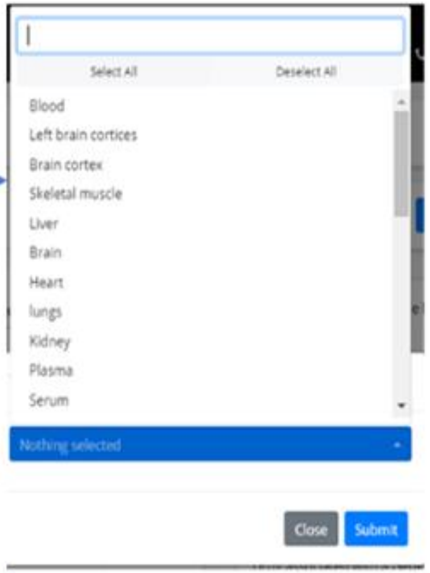
Min : 3000

Max : 4500


Chromosome Location

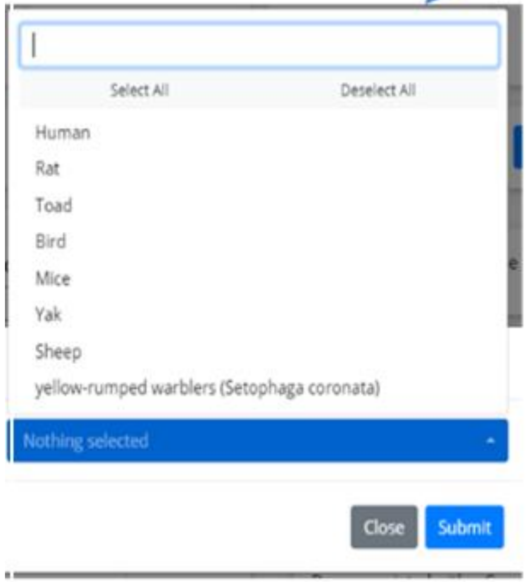

Duration of Experiment

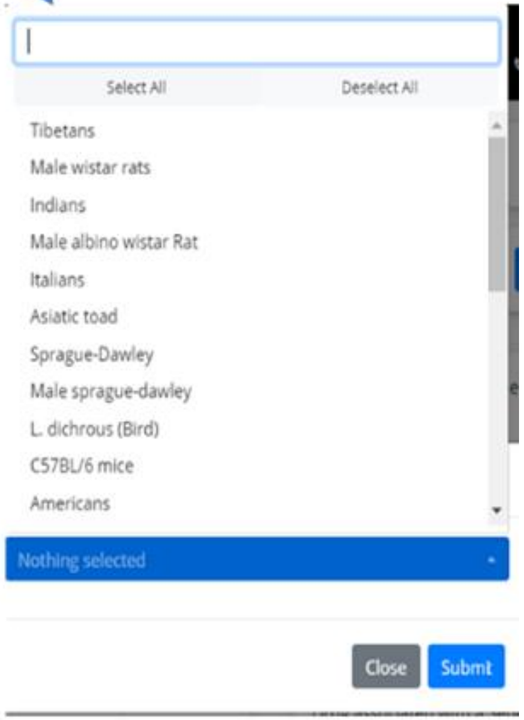
☐ Hours
☐ Days
☐ Weeks
☐ Months
☐ Years


Blood
Left brain cortices
Brain cortex
Skeletal muscle
Liver
Brain
Heart
lungs
Kidney
Plasma
Serum

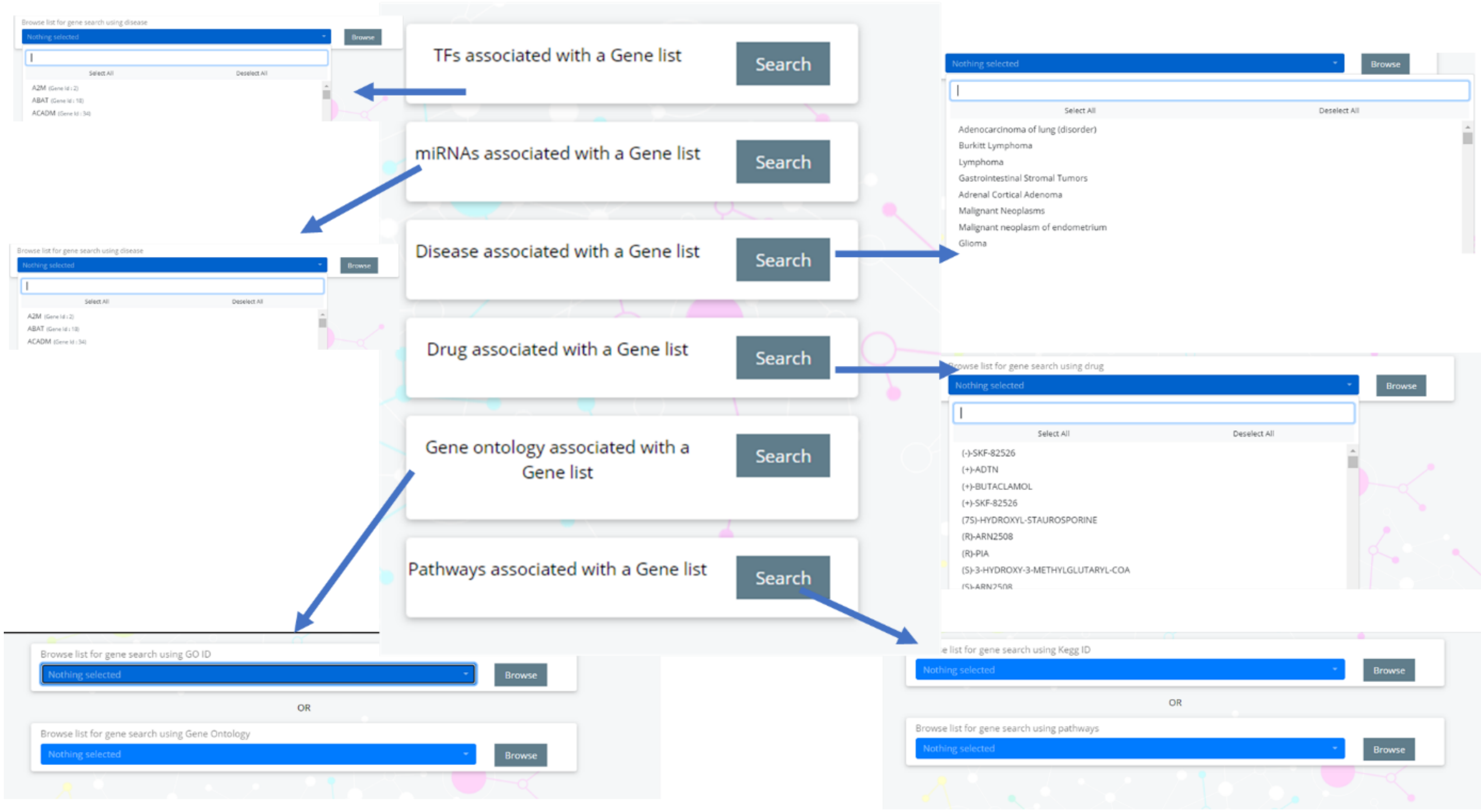
15. First search option is based on single feature search. Specially direct high altitude associated feature such as chromosome location, Duration of experiment, tissue of expression, ethnicity (human studies), source organism, level of regulation , geographical location, biomarker association.

16.All different available option examples have been shown here. Searching from any list would leads to a list proteins that are associated with the search feature shown in adjacent column of table. The filtered table can be downloaded in excel as well as CSV format and can be used in further analysis


Human
Rat
Toad
Bird
Mice
Yak
Sheep
yellow-rumped warblers (Setophaga coronata)


Tibetans
Male wistar rats
Indians
Male albino wistar Rat
Italians
Asiatic toad
Sprague-Dawley
Male sprague-dawley
L. dichrous (Bird)
C57BL/6 mice
Americans

17. In indirect feature option similar results can be seen as in high altitude associated features search. The examples of this feature search shown. Similar type of table would appear in this as in high altitude associated features search. This filtered table can also be downloaded in excel as well as CSV format and can be used in further analysis.



A2M (Gene Id : 2)
ABAT (Gene Id : 18)
ACADM (Gene Id : 34)

A2M (Gene Id : 2)
ABAT (Gene Id : 18)
ACADM (Gene Id : 34)

Adenocarcinoma of lung (disorder)
Burkitt Lymphoma
Lymphoma
Gastrointestinal Stromal Tumors
Adrenal Cortical Adenoma
Malignant Neoplasms
Malignant neoplasm of endometrium
Glioma

(-)-SKF-82526
(+)-ADTN
(+)-BUTACLAMOL
(+)-SKF-82526
(7S)-HYDROXYL-STAUROSPORINE
(R)-ARN2508
(R)-PIA
(S)-3-HYDROXY-3-METHYLGLUTARYL-COA
(S)-ARN2508