Phenotype and genotype input file preparation:

For Phenotype analysis tab

- The phenotype file should contain at least five columns such as Location, Season, Block, Accession, Trait1, Trait2 etc.
  Where,
  Trait1, Trait2 represents phenotype value of various traits and any number of traits can be added in the phenotype file for phenotype analysis.

For 3k Rice (Default)

Location	Season	BLOCK	Accession	NPT	PL	GYKGPHA
ISARC	20WS	1	POHHERLIMASION::IRGC 62025-1	5	28.9	2011.7
ISARC	21WS	1	POHHERLIMASION::IRGC 62025-1	8	35.33	3929.825
ISARC	21WS	1	NAN TEO 14::IRGC 7304-1	10	25.4	3512.069
ISARC	20WS	1	NAN TEO 14::IRGC 7304-1	6	19.6	
ISARC	21WS	1	URAIBOOL::IRGC 52785-1	7.67	28.9	2601.887
ISARC	20WS	1	URAIBOOL::IRGC 52785-1	5	25.8	3729.09
ISARC	20WS	1	IARI 11387::IRGC 19573-1	6	23.8	2347.06
ISARC	21WS	1	IARI 11387::IRGC 19573-1	7	20.53	3091.667
ISARC	21WS	1	HARRA::GERVEX 501-C1	6.67	18.4	749.1228
ISARC	20WS	1	HARRA::GERVEX 501-C1	7	12.5	478.33
ISARC	20WS	1	PL 3165::IRGC 62827-1	6	13.8	630
ISARC	21WS	1	PL 3165::IRGC 62827-1	8.67	18.13	1878.333
ISARC	20WS	1	ARC 18371::IRGC 42423-2	6	26.1	2444.26
ISARC	21WS	1	ARC 18371::IRGC 42423-2	5.67	21.3	3120.339

For other diploid organism (external)

Trial	Genotype	Replication	TN	RN	ARP	RL	subpopulation
2nd Expt	MG1	3	9	212	33.99	711.243	xx
2nd Expt	MG1	2	9	235	40.06	701.678	xx
2nd Expt	MG1	1	10	283	37.82	524.315	xx
2nd Expt	MG5	1	8	163	34.65	732.077	xx
2nd Expt	MG5	2	9	177	37.64	617.693	xx
2nd Expt	MG5	3	9	196	31.9	892.359	xx
2nd Expt	MG6	1	9	218	24.79	828.369	xx
2nd	MG6	3	9	226	24.7	589.501	xx

Expt							
2nd Expt	MG6	2	9	239	25.79	507.984	yy

For other analysis such as Genofile Extract, GWAS, Et-GWAS, Haplopheno tabs

External: Accession/ IDs, should be same

Default

- First column name should be X.Phenotype of phenotype file(.csv format).
- ID's in X.Phenotype. column should be in the following format e.g; IRIS_313-11586.
- No need to provide a subpopulation column.

External

- The first column name should be X.Phenotype of phenotype file.
- Accession/ IDs in X.Phenotype. column should be the same as the vcf file (genotype file) provided by the user.
- Subpopulation or group information should be provided in the third column of the phenotype file and name of the **column should be "subpopulation"**.

For 3k Rice (Default)

| X.Phenotype.	PL
IRIS_313-11586	27.23
IRIS_313-11037	27.83
IRIS_313-8061	19.63
IRIS_313-8060	26.07
IRIS_313-8057	26.83
IRIS_313-9140	24.33
IRIS_313-12185	27.5
IRIS_313-11744	24.73
IRIS_313-12057	29.07

For other diploid organism (external)

| X.Phenotype.	ARP(adjusted)	subpopulation
MG1	37.00889	yq
MG10	29.31557	yy
MG100	30.35416	yy
MG101	30.87122	xx
MG106	23.78305	zz
MG109	21.66455	yy
MG11	34.11961	zz
MG110	35.22003	xx
MG114	31.79448	zz

Run HapGUI from docker hub: (follow video manual)

1. Download docker desktop based on operating system from
   https://www.docker.com/products/docker-desktop/
2. Follow the instructions for installing docker by following the provided video manual.
3. To run HapGUI, users need to follow the following instructions just once.
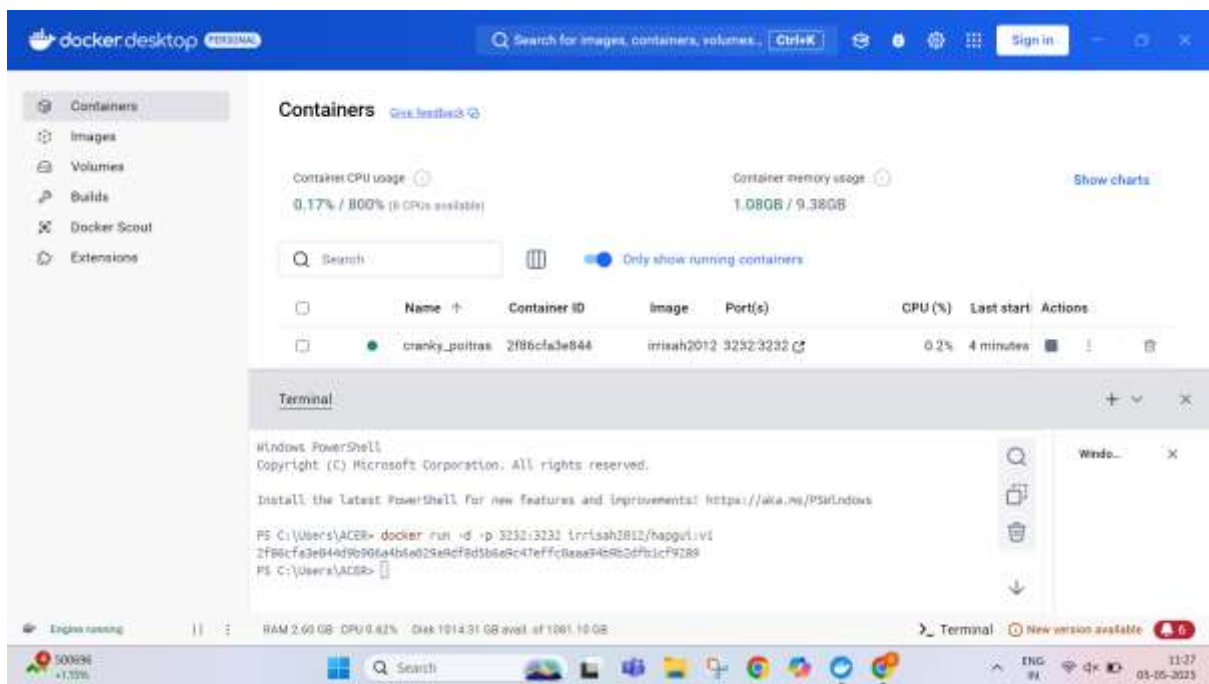   Then open the terminal of docker and type the following commands one by one.

   docker pull irrisah2012/hapgui:v1
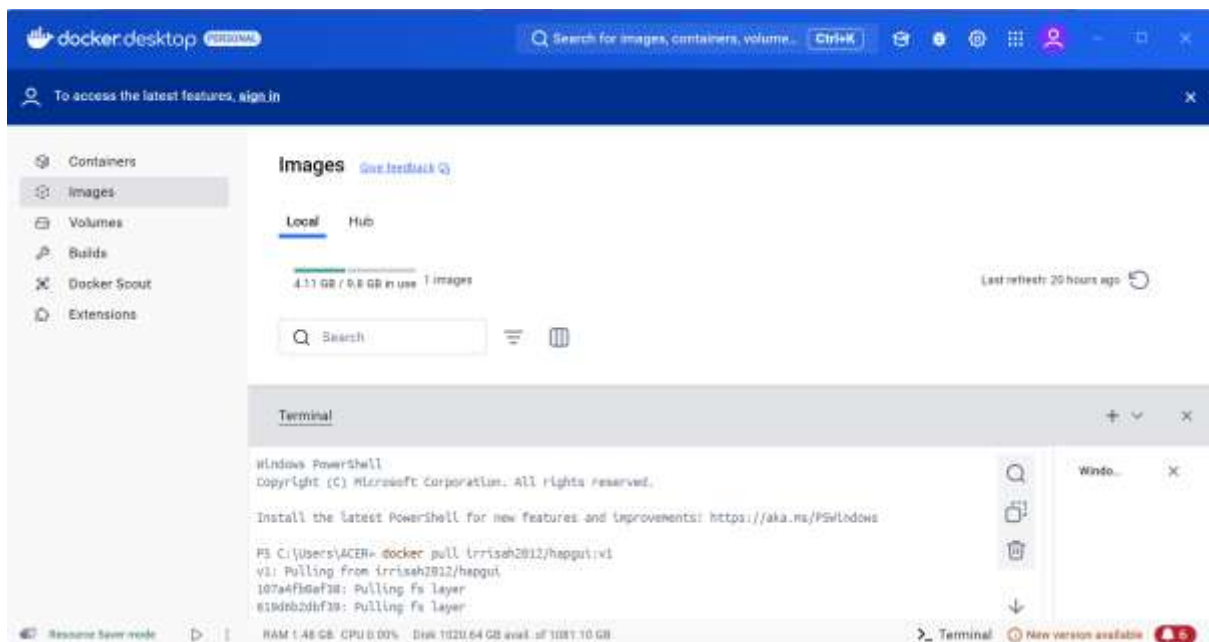
After the pull completes the docker terminal will look like the following.
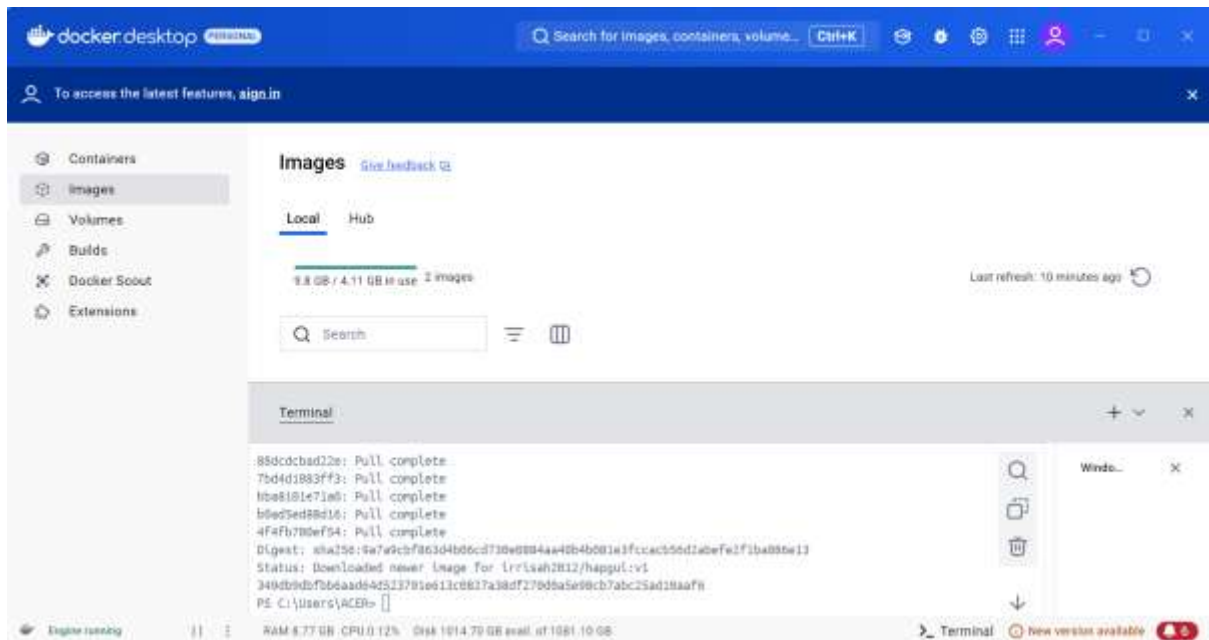
docker run -d -p 3232:3232 irrisah2012/hapgui:v1

4. Then click on the highlighted symbol, after that on the browser (chrome, firefox etc.) of the user's PC HapGUI will open as provided in the screen short below.
    Else, copy the link (http://localhost:3232/) on browser and press enter. Then HapGUI interface will open as shown below.
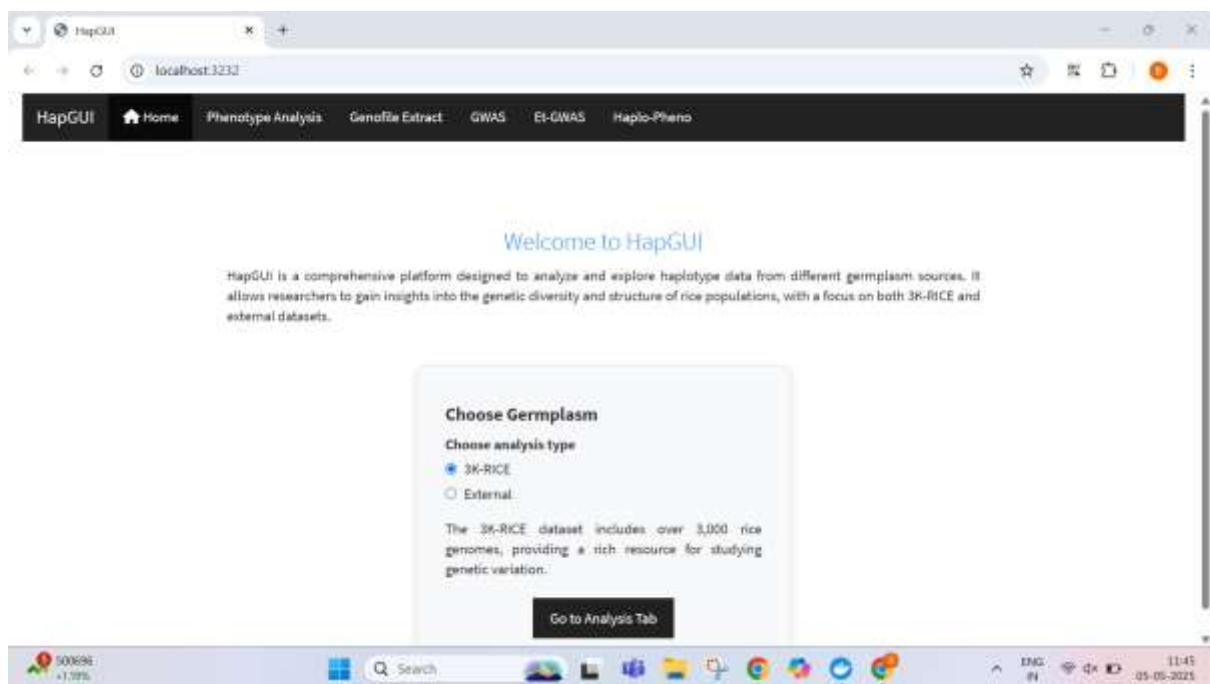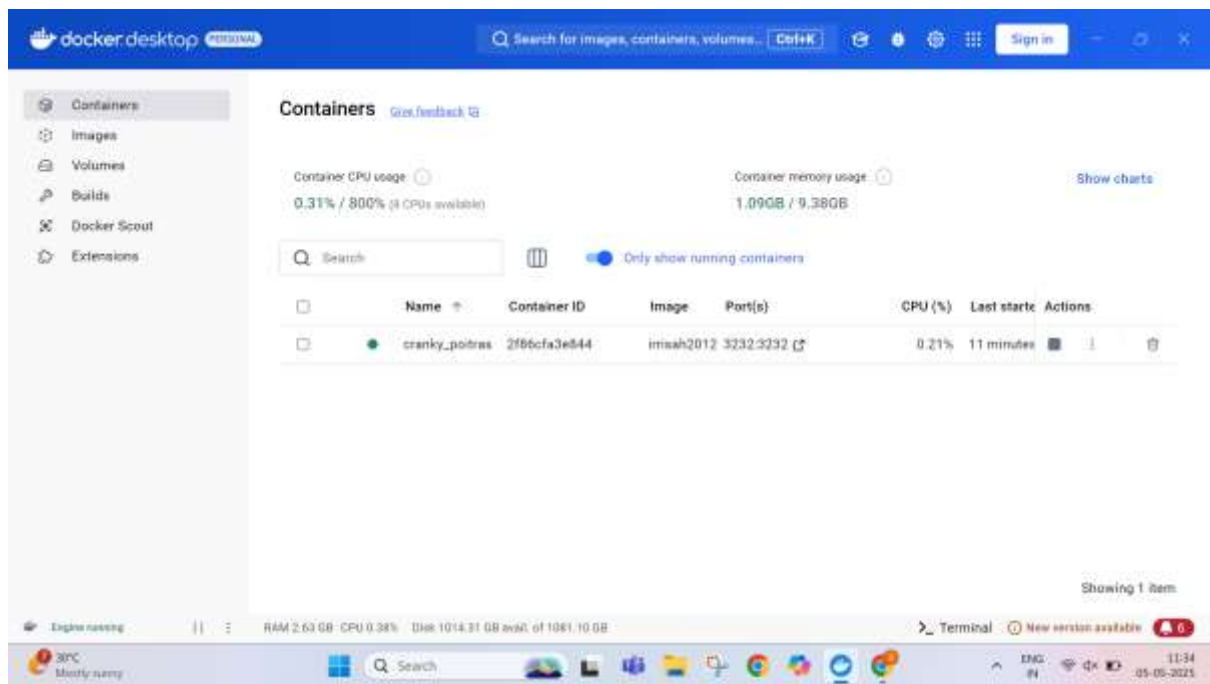
Note: Before performing Genofile Extract (Genotype file extraction) for GWAS, Et-GWAS and Haplopheno analysis for other panels of Rice as well as other diploid organism gff3 file(General Feature Format) need to be downloaded from ensemble database (https://ensemblgenomes.org/) as explained in the video manual.