

The SNPstar manual is NOT completely finished! The complete version is available at paper submission!

Manual SNPstar

Index

1 Use of SNPstar for scientific research.....	3
2 General options.....	3
2.1 Gene selection.....	3
2.1.1 Select a gene and transcript via AT number.....	4
2.1.2 Select a group of genes.....	4
2.2 Accession selection.....	5
2.2.1. Direct selection all or specified accessions.....	5
2.2.2 Select a subset of accessions via uploading accession IDs.....	6
2.2.3 Selecting accession via 1001 genomes groups.....	9
2.2.4 Select a subset of accessions via Koeppen climate region.....	11
2.2.5 Select a subset of accessions via country.....	14
2.3 Compute results.....	17
2.4 Results for selection.....	18
2.4.1 Sequence viewer.....	19
2.4.2 SNPs.....	26
2.4.3 SNPs ↔ Accessions.....	29
2.4.4 DNA Haplotype ↔ Accessions.....	30
2.4.5 Proteotype ↔ Accession.....	35
2.4.6 Download options.....	40
2.4.6.1 Download all.....	40
2.4.6.2 DNA haplotype table.....	42
2.4.6.3 Proteotype sequence complete AA.....	43
2.4.6.4 Haplotype sequences complete CDS.....	43
2.4.6.5 Proteotype table.....	44
2.4.6.6 SNP table.....	44
2.4.6.7 SNP accession table.....	47

2.4.6.8	Accession information.....	48
2.5	Search SNP database.....	48
2.5.1	Get SNP information.....	49
2.5.2	Get accessions for SNPs.....	51
2.6	Build & Download GWAS matrices.....	53
2.6.1	Layout of SNP GWAS Matrix.....	58
2.6.2	Layout of Haplotype GWAS Matrix.....	58
2.6.3	Layout of Proteotype GWAS Matrix.....	59
3	Use cases.....	59
3.1	Single target gene.....	59
3.1.1	Identification of SNPs in a single gene of interest.....	60
3.1.1.1	General selection and SNP identification.....	60
3.1.1.2	Identification of accessions with SNPs of interest.....	64
3.1.1.3	Identification of accessions with haplotypes of interest.....	67
3.1.1.4	Identification of accessions with protein haplotypes of interest.....	70
3.1.2	Identification of known associated SNPs in a gene of interest.....	70
3.2	GWAS.....	71
3.2.1	GWAS with all transcripts and a subset of accession.....	71
3.2.2	GWAS with a subset of transcripts and a subset of accessions.....	71
3.2.3	GWAS with all or a subset of transcripts and all accessions.....	71


1 Use of SNPstar for scientific research

What can SNPstar do for you? SNPstar is a tool to help researchers in different ways. SNPstar serves as some sort of SNP atlas which collects different kinds of information, characteristics and features for each SNP in genic regions based on genomic information from the 1001 genomes project (<https://1001genomes.org/>) of *Arabidopsis thaliana*. Currently, SNPstar has a focus on nonsynonymous SNPs, thereby serving as a starting point to gather meaningful information about SNPs that potentially affect protein structure and function. This manual provides a step-by-step guideline for gathering this information for SNPs in (a) gene(s) of interest based on the complete or a subset of the naturally occurring accessions present in the 1001 genomes collection.


2 General options

On the landing page you initially select your accession set of interest. Here, we will introduce SNPstar on the basis of the 1001 genomes dataset for *Arabidopsis thaliana*.

Welcome to the SNPstar web server!



Free access also for commercial users.

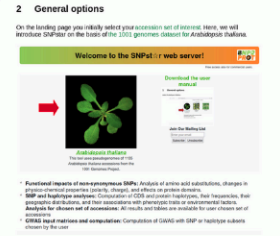


Arabidopsis thaliana
This tool uses pseudogenomes of 1135 *Arabidopsis thaliana* accessions from the 1001 Genomes Project.

Download the user manual

2 General options

On the landing page you initially select your accession set of interest. Here, we will introduce SNPstar on the basis of the 1001 genomes dataset for *Arabidopsis thaliana*.



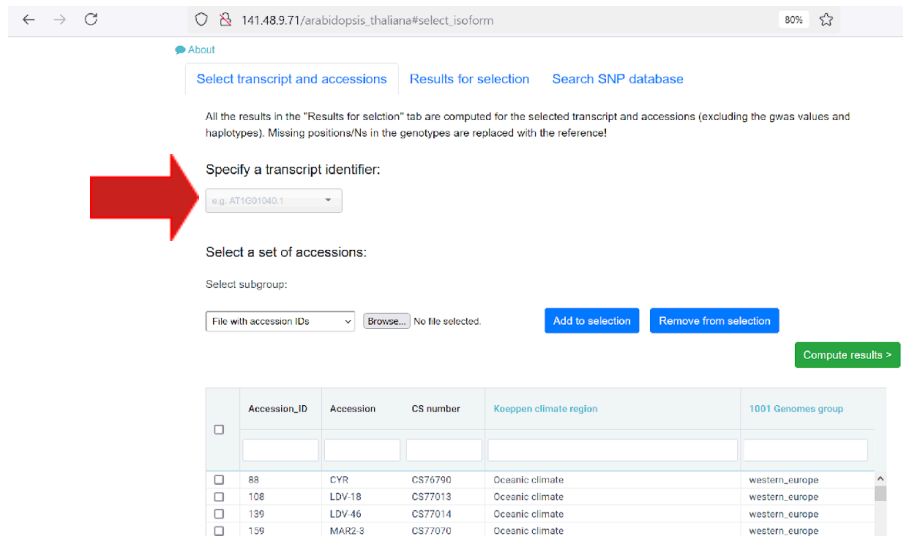
Join Our Mailing List

- ★ **Functional impacts of non-synonymous SNPs:** Analysis of amino acid substitutions, changes in physico-chemical properties (polarity, charge), and effects on protein domains.
- ★ **SNP and haplotype analyses:** Computation of CDS and protein haplotypes, their frequencies, their geographic distributions, and their associations with phenotypic traits or environmental factors.
- ★ **Analysis for chosen set of accessions:** All results and tables are available for sets of accessions chosen by the user
- ★ **GWAS input matrices and computation:** Computation of GWAS with SNP or haplotype subsets chosen by the user

2.1 Gene selection

2.1.1 Select a gene and transcript via AT number

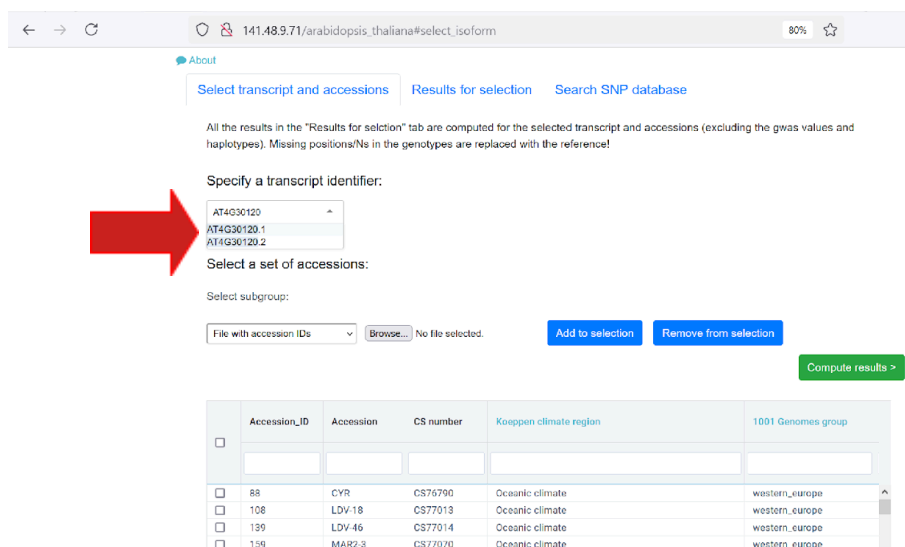
The second step is to select your gene of interest. As an example, we will introduce all the features of SNPstar on the basis of the gene *HMA3* (AT4G30120). You can select any gene of interest either by typing the common gene name (e.g. HMA3) or the AGI identifier in the drop down menu as shown below.



The screenshot shows the SNPstar web interface. The browser address bar displays '141.48.9.71/arabidopsis_thaliana#select_isoform'. The page has three tabs: 'Select transcript and accessions', 'Results for selection', and 'Search SNP database'. Below the tabs, there is a note: 'All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!'. The 'Specify a transcript identifier:' section has a dropdown menu with 'e.g. AT1G01040.1' selected. A red arrow points to the dropdown menu. Below this, the 'Select a set of accessions:' section has a 'File with accession IDs' dropdown, a 'Browse...' button, and 'No file selected.' text. There are 'Add to selection' and 'Remove from selection' buttons. A 'Compute results >' button is at the bottom right. A table is displayed below the form with the following data:

	Accession_ID	Accession	CS number	Koepfen climate region	1001 Genomes group
<input type="checkbox"/>					
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe

Select your gene model (transcript) of interest. Here, we select AT4G30120.1, but you can select any of the other known splice variants.



The screenshot shows the SNPstar web interface. The browser address bar displays '141.48.9.71/arabidopsis_thaliana#select_isoform'. The page has three tabs: 'Select transcript and accessions', 'Results for selection', and 'Search SNP database'. Below the tabs, there is a note: 'All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!'. The 'Specify a transcript identifier:' section has a dropdown menu with 'AT4G30120.1' selected. A red arrow points to the dropdown menu. Below this, the 'Select a set of accessions:' section has a 'File with accession IDs' dropdown, a 'Browse...' button, and 'No file selected.' text. There are 'Add to selection' and 'Remove from selection' buttons. A 'Compute results >' button is at the bottom right. A table is displayed below the form with the following data:

	Accession_ID	Accession	CS number	Koepfen climate region	1001 Genomes group
<input type="checkbox"/>					
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe

2.1.2 Select a group of genes

In progress!

2.2 Accession selection

There are several options to select subsets of accessions for your analysis. We will introduce these in the following subchapters 2.2.1 - 2.2.5.

2.2.1. Direct selection all or specified accessions

You can select either the complete set of 1135 accessions or, alternatively, compile a subpopulation by checking the boxes on the left.

141.48.9.71/arabidopsis_thaliana#select_isoform

Select transcript and accessions | Results for selection | Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:
AT4G30120.1

Select a set of accessions:
Select subgroup:
File with accession IDs | Browse... No file selected. | Add to selection | Remove from selection

Compute results >

	Accession_ID	Accession	CS number	Koepfen climate region	1001 Genomes group
<input checked="" type="checkbox"/>					
<input checked="" type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	199	LDV-46	CS77014	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	199	MAR2-3	CS77070	Oceanic climate	western_europe

141.48.9.71/arabidopsis_thaliana#select_isoform

Select transcript and accessions | Results for selection | Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:
AT4G30120.1

Select a set of accessions:
Select subgroup:
File with accession IDs | Browse... No file selected. | Add to selection | Remove from selection

Compute results >

	Accession_ID	Accession	CS number	Koepfen climate region	1001 Genomes group
<input checked="" type="checkbox"/>					
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	199	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	199	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PYL-6	CS77198	Oceanic climate	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe

To finish the selection click 'Add to selection'.

← → ↻ 141.48.9.71/arabidopsis_thaliana#select_isoform 80% ☆

About

Select transcript and accessions Results for selection Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:
 AT4G30120.1

Select a set of accessions:
 Select subgroup:
 Complete 1135 accessions Add to selection Remove from selection

Compute results >

<input type="checkbox"/>	Accession_ID	Accession	CS number	Koepfen climate region	1001 Genomes group
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe

You can remove the selection by clicking 'Remove from selection'.

← → ↻ 141.48.9.71/arabidopsis_thaliana#select_isoform 80% ☆

About

Select transcript and accessions Results for selection Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:
 AT4G30120.1

Select a set of accessions:
 Select subgroup:
 File with accession IDs Browse... No file selected. Add to selection Remove from selection

Compute results >

<input checked="" type="checkbox"/>	Accession_ID	Accession	CS number	Koepfen climate region	1001 Genomes group
<input checked="" type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe

2.2.2 Select a subset of accessions via uploading accession IDs

Instead of manually checking all the corresponding boxes, you can upload a file of accessions by marking 'File with accession IDs' in the drop down menu.

About

Select transcript and accessions Results for selection Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:

AT4G30120.1

Select a set of accessions:

Select subgroup:

File with accession IDs No file selected.

File with accession IDs
 1001 Genomes Group
 Koeppen climate region
 Country
 Complete 1135 accessions

	Accession	CS number	Koeppen climate region	1001 Genomes group	
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PYL-6	CS77198	Oceanic climate	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input type="checkbox"/>	410	Doubravnik7	CS76808	Oceanic climate	central_europe
<input type="checkbox"/>	424	Draha2	CS76812	Oceanic climate	central_europe

The file with the accession IDs should be in the following format:

	A	B
1	88	
2	108	
3	159	
4	350	
5	351	
6	410	
7		
8		
9		
10		

Click 'Browse' and upload your specified accession list.

About

Select transcript and accessions Results for selection Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:
 AT4G30120.1

Select a set of accession:
 Select subgroup:
 File with accession IDs Browse... No file selected. Add to selection Remove from selection

Compute results >

	Accession_ID	Accession	CS number	Koepfen climate region	1001 Genomes group
<input type="checkbox"/>					
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PYL-6	CS77198	Oceanic climate	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input type="checkbox"/>	410	Doubravnik7	CS76808	Oceanic climate	central_europe

Click 'Add to selection' for adding.

About

Select transcript and accessions Results for selection Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:
 AT4G30120.1

Select a set of accessions:
 Select subgroup:
 File with accession IDs Browse... SNPstar_accessions.csv Add to selection Remove from selection

Compute results >

	Accession_ID	Accession	CS number	Koepfen climate region	1001 Genomes group
<input type="checkbox"/>					
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PYL-6	CS77198	Oceanic climate	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input type="checkbox"/>	410	Doubravnik7	CS76808	Oceanic climate	central_europe

About

Select transcript and accessions Results for selection Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:

AT4G30120.1

Select a set of accessions:

Select subgroup:

File with accession IDs Browse... SNPstar_accessions.csv Add to selection Remove from selection

Compute results >

	Accession_ID	Accession	CS number	Koeppen climate region	1001 Genomes group
<input type="checkbox"/>					
<input checked="" type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PYL-6	CS77198	Oceanic climate	admixed
<input checked="" type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input checked="" type="checkbox"/>	410	Doubravnik7	CS76808	Oceanic climate	central_europe

You can always remove the selection via 'Remove from selection' (see chapter 2.2.1).

2.2.3 Selecting accession via 1001 genomes groups

Another straight-forward option to select sub-populations is to select genetic diversity groups from the 1001 genomes project. Use the drop-down menu and select '1001 Genomes Group'.

About

Select transcript and accessions Results for selection Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:

AT4G30120.1

Select a set of accessions:

Select subgroup:

File with accession IDs Browse... SNPstar_accessions.csv Add to selection Remove from selection

1001 Genomes Group

Compute results >

	Accession_ID	Accession	CS number	Koeppen climate region	1001 Genomes group
<input type="checkbox"/>					
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PYL-6	CS77198	Oceanic climate	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input type="checkbox"/>	410	Doubravnik7	CS76808	Oceanic climate	central_europe

Next, a second drop-down menu appears where you can select the genetic diversity group.

Manual About

Select transcript and accessions **Results for selection** Search SNP database Build & Download GWAS matrices **SNP2PROT**

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/ Ns in the genotypes are replaced with the reference

Specify a transcript identifier or gene name:
 AT4G3120.1

Select a set of accessions:
 Select subgroup:
 1001 Genomes Group Admixed **Add to selection** Remove from selection Compute results >

Accession_ID (1001 genomes)	Accession	CS number	Koepfen climate region	1001 Genomes group	
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PYL-6	CS77198	Oceanic climate	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input type="checkbox"/>	410	Doubraunik7	CS76808	Oceanic climate	central_europe
<input type="checkbox"/>	424	Draha2	CS76812	Oceanic climate	central_europe
<input type="checkbox"/>	428	Borky1	CS76718	Oceanic climate	central_europe
<input type="checkbox"/>	430	Gr-1	CS76496	Warm summer continental or hemiboreal climates	central_europe
<input type="checkbox"/>	470	BRR4	CS78943	Hot summer continental climates	germany
<input type="checkbox"/>	476	BRR12	CS78944	Hot summer continental climates	germany
<input type="checkbox"/>	484	BRR23	CS78945	Hot summer continental climates	germany
<input type="checkbox"/>	504	BRR57	CS78946	Hot summer continental climates	germany
<input type="checkbox"/>	506	BRR60	CS78947	Hot summer continental climates	germany
<input type="checkbox"/>	531	BRR107	CS78948	Hot summer continental climates	germany

Manual About

Select transcript and accessions **Results for selection** Search SNP database Build & Download GWAS matrices **SNP2PROT**

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/ Ns in the genotypes are replaced with the reference

Specify a transcript identifier or gene name:
 AT4G3120.1

Select a set of accessions:
 Select subgroup:
 1001 Genomes Group Admixed **Add to selection** Remove from selection Compute results >

Admixed
 Asia
 Central europe
 Germany
 Italy, Balkan, Caucasus
 North sweden
 Relict
 South sweden
 Spain
 Western europe

Accession_ID (1001 genomes)	Accession	CS number	Koepfen climate region	1001 Genomes group	
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PYL-6	CS77198	Oceanic climate	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input type="checkbox"/>	410	Doubraunik7	CS76808	Oceanic climate	central_europe
<input type="checkbox"/>	424	Draha2	CS76812	Oceanic climate	central_europe
<input type="checkbox"/>	428	Borky1	CS76718	Oceanic climate	central_europe
<input type="checkbox"/>	430	Gr-1	CS76496	Warm summer continental or hemiboreal climates	central_europe
<input type="checkbox"/>	470	BRR4	CS78943	Hot summer continental climates	germany
<input type="checkbox"/>	476	BRR12	CS78944	Hot summer continental climates	germany
<input type="checkbox"/>	484	BRR23	CS78945	Hot summer continental climates	germany
<input type="checkbox"/>	504	BRR57	CS78946	Hot summer continental climates	germany
<input type="checkbox"/>	506	BRR60	CS78947	Hot summer continental climates	germany
<input type="checkbox"/>	531	BRR107	CS78948	Hot summer continental climates	germany

Click 'Add to selection'.

Select transcript and accessions **Results for selection** Search SNP database Build & Download GWAS matrices **SNP2PROT** Manual About

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/ Ns in the genotypes are replaced with the reference

Specify a transcript identifier or gene name:

Select a set of accessions:
 Select subgroup:

Accession_ID (1001 genomes)	Accession	CS number	Koeppen climate region	1001 Genomes group	
<input type="checkbox"/>					
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	139	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PVL-6	CS77198	Oceanic climate	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input type="checkbox"/>	410	Doubravnik7	CS76808	Oceanic climate	central_europe
<input type="checkbox"/>	424	Draha2	CS76812	Oceanic climate	central_europe
<input type="checkbox"/>	428	Borky1	CS76718	Oceanic climate	central_europe
<input type="checkbox"/>	430	Gr-1	CS76496	Warm summer continental or hemiboreal climates	central_europe
<input type="checkbox"/>	470	BRR4	CS78943	Hot summer continental climates	germany
<input type="checkbox"/>	476	BRR12	CS78944	Hot summer continental climates	germany
<input type="checkbox"/>	484	BRR23	CS78945	Hot summer continental climates	germany
<input type="checkbox"/>	504	BRR57	CS78946	Hot summer continental climates	germany
<input type="checkbox"/>	506	BRR60	CS78947	Hot summer continental climates	germany
<input type="checkbox"/>	531	BRR107	CS78948	Hot summer continental climates	germany

Select transcript and accessions **Results for selection** Search SNP database Build & Download GWAS matrices **SNP2PROT** Manual About

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/ Ns in the genotypes are replaced with the reference

Specify a transcript identifier or gene name:

Select a set of accessions:
 Select subgroup:

Accession_ID (1001 genomes)	Accession	CS number	Koeppen climate region	1001 Genomes group	
<input type="checkbox"/>					
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	139	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PVL-6	CS77198	Oceanic climate	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input checked="" type="checkbox"/>	410	Doubravnik7	CS76808	Oceanic climate	central_europe
<input checked="" type="checkbox"/>	424	Draha2	CS76812	Oceanic climate	central_europe
<input checked="" type="checkbox"/>	428	Borky1	CS76718	Oceanic climate	central_europe
<input checked="" type="checkbox"/>	430	Gr-1	CS76496	Warm summer continental or hemiboreal climates	central_europe
<input type="checkbox"/>	470	BRR4	CS78943	Hot summer continental climates	germany
<input type="checkbox"/>	476	BRR12	CS78944	Hot summer continental climates	germany
<input type="checkbox"/>	484	BRR23	CS78945	Hot summer continental climates	germany
<input type="checkbox"/>	504	BRR57	CS78946	Hot summer continental climates	germany
<input type="checkbox"/>	506	BRR60	CS78947	Hot summer continental climates	germany
<input type="checkbox"/>	531	BRR107	CS78948	Hot summer continental climates	germany

It's possible to add more 1001 Genome groups if you repeat the procedure with other regions of origin. You can remove specific groups via 'Remove from selection' (see chapter 2.2.1).

2.2.4 Select a subset of accessions via Koeppen climate region

If you would like to select a subset of available accession based on the Koeppen climate regions then use the drop-down menu Select subgroup and select Koeppen climate region. The climate regions according to the Koeppen climate classifications.

[About](#)
[Select transcript and accessions](#)
[Results for selection](#)
[Search SNP database](#)

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:

AT4G30120.1

Select a set of accessions:

Select subgroup:

File with accession IDs SNPstar_accessions.csv

File with accession IDs

- 1001 Genomes Group
- Koepppen climate region**
- Country
- Complete 1135 accessions

	Accession	CS number	Koepppen climate region	1001 Genomes group	
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PYL-6	CS77198	Oceanic climate	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input type="checkbox"/>	410	Doubravnik7	CS76808	Oceanic climate	central_europe
<input type="checkbox"/>	424	Draha2	CS76812	Oceanic climate	central_europe

A second drop-down menu appears. In that menu you can select your climate regions of interest.

[Manual](#)
[About](#)
[Select transcript and accessions](#)
[Results for selection](#)
[Search SNP database](#)
[Build & Download GWAS matrices](#)

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier or gene name:

AT4G30120.1

Select a set of accessions:

Select subgroup:

Koepppen climate region

	Accession_ID (1001 genomes)	Accession	CS number	Koepppen climate region	1001 Genomes group
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PYL-6	CS77198	Oceanic climate	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input type="checkbox"/>	410	Doubravnik7	CS76808	Oceanic climate	central_europe
<input type="checkbox"/>	424	Draha2	CS76812	Oceanic climate	central_europe
<input type="checkbox"/>	428	Borky1	CS76718	Oceanic climate	central_europe
<input type="checkbox"/>	430	Gr-1	CS76496	Warm summer continental or hemiboreal climates	central_europe
<input type="checkbox"/>	470	BRR4	CS78943	Hot summer continental climates	germany
<input type="checkbox"/>	476	BRR12	CS78944	Hot summer continental climates	germany
<input type="checkbox"/>	484	BRR23	CS78945	Hot summer continental climates	germany
<input type="checkbox"/>	504	BRR57	CS78946	Hot summer continental climates	germany
<input type="checkbox"/>	506	BRR60	CS78947	Hot summer continental climates	germany
<input type="checkbox"/>	531	BRR107	CS78948	Hot summer continental climates	germany

Now you can select your 1001 Genome groups of interest (e.g., Oceanic climate, etc.).

About

Select transcript and accessions Results for selection Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:
 AT4G30120.1

Select a set of accessions:
 Select subgroup:
 Koeppen climate region

Subarctic or boreal Add to selection Remove from selection

Oceanic climate
 Warm summer continental or hemiboreal climates
 Hot summer continental climates
 Humid subtropical climates
 Subarctic or boreal climates
 Tundra climate
 Subpolar oceanic climate
 Arid climate
 Mediterranean hot summer climates
 Mediterranean warm/cool summer climates
 NA
 Semi-arid (steppe) climate

Compute results >

Accession_ID	Accession	CS number	Koeppen climate region	1001 Genomes group
<input type="checkbox"/>				
<input type="checkbox"/>	88	CYR		
<input type="checkbox"/>	108	LDV-1		western_europe
<input type="checkbox"/>	139	LDV-4		western_europe
<input type="checkbox"/>	159	MAR2		western_europe
<input type="checkbox"/>	265	PYL-6		admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	central_europe
<input type="checkbox"/>	410	Doubravnik7	CS76808	central_europe
<input type="checkbox"/>	424	Draha2	CS76812	central_europe
<input type="checkbox"/>	428	Borky1	CS76718	central_europe
<input type="checkbox"/>	430	Gr-1	CS76496	central_europe
<input type="checkbox"/>	470	BRR4	CS78943	germany
<input type="checkbox"/>	476	BRR12	CS78944	germany
<input type="checkbox"/>	484	BRR23	CS78945	germany
<input type="checkbox"/>	504	BRR57	CS78946	germany
<input type="checkbox"/>	506	BRR60	CS78947	germany
<input type="checkbox"/>	531	BRR107	CS78948	germany

Compute results >

Click 'Add to selections'.

About

Select transcript and accessions Results for selection Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:
 AT4G30120.1

Select a set of accessions:
 Select subgroup:
 Koeppen climate region Oceanic climate

Add to selection Remove from selection

Compute results >

Accession_ID	Accession	CS number	Koeppen climate region	1001 Genomes group
<input type="checkbox"/>				
<input type="checkbox"/>	88	CYR	CS76790	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	western_europe
<input type="checkbox"/>	265	PYL-6	CS77198	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	central_europe
<input type="checkbox"/>	410	Doubravnik7	CS76808	central_europe
<input type="checkbox"/>	424	Draha2	CS76812	central_europe
<input type="checkbox"/>	428	Borky1	CS76718	central_europe
<input type="checkbox"/>	430	Gr-1	CS76496	central_europe
<input type="checkbox"/>	470	BRR4	CS78943	germany
<input type="checkbox"/>	476	BRR12	CS78944	germany
<input type="checkbox"/>	484	BRR23	CS78945	germany
<input type="checkbox"/>	504	BRR57	CS78946	germany
<input type="checkbox"/>	506	BRR60	CS78947	germany
<input type="checkbox"/>	531	BRR107	CS78948	germany

Compute results >

[About](#) [Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#)

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference!

Specify a transcript identifier:

AT4G30120.1

Select a set of accessions:

Select subgroup: Koeppen climate region Oceanic climate [Add to selection](#) [Remove from selection](#)

[Compute results >](#)

<input type="checkbox"/>	Accession_ID	Accession	CS number	Koeppen climate region	1001 Genomes group
<input checked="" type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	265	PYL-6	CS77198	Oceanic climate	admixed
<input checked="" type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input checked="" type="checkbox"/>	410	Doubravnik7	CS76808	Oceanic climate	central_europe
<input checked="" type="checkbox"/>	424	Draha2	CS76812	Oceanic climate	central_europe
<input checked="" type="checkbox"/>	428	Borky1	CS76718	Oceanic climate	central_europe
<input type="checkbox"/>	430	Gr-1	CS76496	Warm summer continental or hemiboreal climates	central_europe
<input type="checkbox"/>	470	BRR4	CS78943	Hot summer continental climates	germany
<input type="checkbox"/>	476	BRR12	CS78944	Hot summer continental climates	germany
<input type="checkbox"/>	484	BRR23	CS78945	Hot summer continental climates	germany
<input type="checkbox"/>	504	BRR57	CS78946	Hot summer continental climates	germany
<input type="checkbox"/>	506	BRR60	CS78947	Hot summer continental climates	germany
<input type="checkbox"/>	531	BRR107	CS78948	Hot summer continental climates	germany

[Compute results >](#)

It's possible to select more than one Koeppen climate region. To do this, repeat the procedure with other climate regions. You can remove specific regions if you select the Koeppen climate region of interest and click 'Remove the selection'.

2.2.5 Select a subset of accessions via country

If you would like to select a subset of available accessions based on country of origin, then use the drop-down menu 'Select subgroup' and select 'Country'.

About

Select transcript and accessions Results for selection Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference

Specify a transcript identifier:

AT4G30120.1

Select a set of accessions:

Select subgroup:

Koepfen climate region Hot summer cor Add to selection Remove from selection

File with accession IDs
1001 Genomes Group
Koeppen climate region
Country
Complete 1135 accessions

Compute results >

	CS number	Koeppen climate region	1001 Genomes group		
<input type="checkbox"/>	88	CYR	CS76790	Oceanic climate	western_europe
<input type="checkbox"/>	108	LDV-18	CS77013	Oceanic climate	western_europe
<input type="checkbox"/>	139	LDV-46	CS77014	Oceanic climate	western_europe
<input type="checkbox"/>	159	MAR2-3	CS77070	Oceanic climate	western_europe
<input type="checkbox"/>	265	PYL-6	CS77198	Oceanic climate	admixed
<input type="checkbox"/>	350	TOU-A1-88	CS77382	Oceanic climate	western_europe
<input type="checkbox"/>	351	TOU-A1-89	CS77383	Oceanic climate	western_europe
<input type="checkbox"/>	403	Zdarec3	CS78873	Oceanic climate	central_europe
<input type="checkbox"/>	410	Doubravník7	CS78808	Oceanic climate	central_europe
<input type="checkbox"/>	424	Draha2	CS76812	Oceanic climate	central_europe
<input type="checkbox"/>	428	Borky1	CS76718	Oceanic climate	central_europe
<input type="checkbox"/>	430	Gr-1	CS76496	Warm summer continental or hemiboreal climates	central_europe
<input type="checkbox"/>	470	BRR4	CS78943	Hot summer continental climates	germany
<input type="checkbox"/>	476	BRR12	CS78944	Hot summer continental climates	germany
<input type="checkbox"/>	484	BRR23	CS78945	Hot summer continental climates	germany
<input type="checkbox"/>	504	BRR57	CS78946	Hot summer continental climates	germany
<input type="checkbox"/>	506	BRR60	CS78947	Hot summer continental climates	germany
<input type="checkbox"/>	531	BRR107	CS78948	Hot summer continental climates	germany

Compute results >



Now you can select your 'Country' of interest (e.g., Sweden, etc.).

About

Select transcript and accessions Results for selection Search SNP database

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference

Specify a transcript identifier:

AT4G30120.1

Select a set of accessions:

Select subgroup:

Country Sweden Add to selection Remove from selection

Georgia
Germany
Greece
India
Ireland
Iran
Japan
Kazakhstan
Kyrgyzstan
Lebanon
Lithuania
Morocco
Netherlands
Norway
Poland
Portugal
Romania
Russian Federation
Serbia
Switzerland
Slovakia
Sweden
Tajikistan
Turkey
Unighted Kingdom
Ukraine
Kosovo
United States of America

Compute results >

Accession_ID	Accession	CS number	Koeppen climate region	1001 Genomes group
<input type="checkbox"/>	9451	Spro 2	nic climate	south_sweden
<input type="checkbox"/>	9452	Spro 3	nic climate	south_sweden
<input type="checkbox"/>	9453	Stenk	nic climate	south_sweden
<input type="checkbox"/>	9454	Stenk	nic climate	south_sweden
<input type="checkbox"/>	9455	Stenk	nic climate	south_sweden
<input type="checkbox"/>	9470	Tur-4	nic climate	south_sweden
<input type="checkbox"/>	9471	Ull-A-1	nic climate	south_sweden
<input type="checkbox"/>	9476	VårA 1	nic climate	germany
<input type="checkbox"/>	9481	Yst-1	nic climate	south_sweden
<input type="checkbox"/>	9503	11C1	nic climate	western_europe
<input type="checkbox"/>	9506	IP-All	terranean hot summer climates	admixed
<input type="checkbox"/>	9507	IP-Cos	terranean hot summer climates	spain
<input type="checkbox"/>	9508	IP-Mo	terranean hot summer climates	admixed
<input type="checkbox"/>	9509	IP-Mo	terranean hot summer climates	spain
<input type="checkbox"/>	9510	IP-Mo	terranean hot summer climates	spain
<input type="checkbox"/>	9511	IP-Mo	terranean hot summer climates	spain
<input type="checkbox"/>	9512	IP-Mo	terranean hot summer climates	spain
<input type="checkbox"/>	9513	IP-Add	terranean hot summer climates	admixed

Compute results >



Click 'Add to selections'.

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference

Specify a transcript identifier:

AT4G30120.1

Select a set of accessions:

Select subgroup:

Country Sweden



Add to selection Remove from selection

Compute results >

Accession_ID	Accession	CS number	Koepfen climate region	1001 Genomes group	
<input type="checkbox"/>					
<input type="checkbox"/>	9451	Spro 2	CS77264	Oceanic climate	south_sweden
<input type="checkbox"/>	9452	Spro 3	CS77265	Oceanic climate	south_sweden
<input type="checkbox"/>	9453	Stenk-2	CS77274	Oceanic climate	south_sweden
<input type="checkbox"/>	9454	Stenk-3	CS77275	Oceanic climate	south_sweden
<input type="checkbox"/>	9455	Stenk-4	CS77276	Oceanic climate	south_sweden
<input type="checkbox"/>	9470	Tur-4	CS77399	Warm summer continental or hemiboreal climates	south_sweden
<input type="checkbox"/>	9471	Ull-A-1	CS78820	Oceanic climate	south_sweden
<input type="checkbox"/>	9476	VårA 1	CS78832	Oceanic climate	germany
<input type="checkbox"/>	9481	Yst-1	CS78869	Oceanic climate	south_sweden
<input type="checkbox"/>	9503	11C1	CS76640	Oceanic climate	western_europe
<input type="checkbox"/>	9506	IP-Alo-0	CS76662	Mediterranean hot summer climates	admixed
<input type="checkbox"/>	9507	IP-Coa-0	CS76775	Mediterranean hot summer climates	spain
<input type="checkbox"/>	9508	IP-Mos-1	CS77108	Mediterranean hot summer climates	admixed
<input type="checkbox"/>	9509	IP-Reg-0	CS77207	Mediterranean hot summer climates	spain
<input type="checkbox"/>	9510	IP-Rei-0	CS77208	Mediterranean hot summer climates	spain
<input type="checkbox"/>	9511	IP-Vav-0	CS78835	Mediterranean hot summer climates	spain
<input type="checkbox"/>	9512	IP-Vid-1	CS78842	Mediterranean hot summer climates	spain
<input type="checkbox"/>	9513	IP-Adc-5	CS76646	Mediterranean hot summer climates	admixed

Compute results >

All the results in the "Results for selection" tab are computed for the selected transcript and accessions (excluding the gwas values and haplotypes). Missing positions/Ns in the genotypes are replaced with the reference

Specify a transcript identifier:

AT4G30120.1

Select a set of accessions:

Select subgroup:

Country Sweden

Add to selection Remove from selection

Compute results >

Accession_ID	Accession	CS number	Koepfen climate region	1001 Genomes group	
<input type="checkbox"/>					
<input checked="" type="checkbox"/>	9451	Spro 2	CS77264	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9452	Spro 3	CS77265	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9453	Stenk-2	CS77274	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9454	Stenk-3	CS77275	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9455	Stenk-4	CS77276	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9470	Tur-4	CS77399	Warm summer continental or hemiboreal climates	south_sweden
<input checked="" type="checkbox"/>	9471	Ull-A-1	CS78820	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9476	VårA 1	CS78832	Oceanic climate	germany
<input checked="" type="checkbox"/>	9481	Yst-1	CS78869	Oceanic climate	south_sweden
<input type="checkbox"/>	9503	11C1	CS76640	Oceanic climate	western_europe
<input type="checkbox"/>	9506	IP-Alo-0	CS76662	Mediterranean hot summer climates	admixed
<input type="checkbox"/>	9507	IP-Coa-0	CS76775	Mediterranean hot summer climates	spain
<input type="checkbox"/>	9508	IP-Mos-1	CS77108	Mediterranean hot summer climates	admixed
<input type="checkbox"/>	9509	IP-Reg-0	CS77207	Mediterranean hot summer climates	spain
<input type="checkbox"/>	9510	IP-Rei-0	CS77208	Mediterranean hot summer climates	spain
<input type="checkbox"/>	9511	IP-Vav-0	CS78835	Mediterranean hot summer climates	spain
<input type="checkbox"/>	9512	IP-Vid-1	CS78842	Mediterranean hot summer climates	spain
<input type="checkbox"/>	9513	IP-Adc-5	CS76646	Mediterranean hot summer climates	admixed

Compute results >

It's possible to add more 'Countries' when you repeat the procedure with other 'countries'. You can remove specific regions if you select the 'Country' of interest and click 'Remove the selection' (see chapter 2.2.1).

2.3 Compute results

To calculate the result of your selections press 'Compute results'.

Results for selection

Specify a transcript identifier:
AT4G30120.1


Select a set of accessions:
Complete 1135 accessions

Add to selection Remove from selection

Compute results >

	Accession_ID	Accession	CS number	Koeppen climate region	1001 Genomes group
<input checked="" type="checkbox"/>					
<input checked="" type="checkbox"/>	9451	Spro 2	CS77264	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9452	Spro 3	CS77265	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9453	Stenk-2	CS77274	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9454	Stenk-3	CS77275	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9455	Stenk-4	CS77276	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9470	Tur-4	CS77399	Warm summer continental or hemiboreal climates	south_sweden
<input checked="" type="checkbox"/>	9471	Ull-A-1	CS78820	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9476	VårA 1	CS78832	Oceanic climate	germany
<input checked="" type="checkbox"/>	9481	Yst-1	CS78869	Oceanic climate	south_sweden
<input checked="" type="checkbox"/>	9503	11C1	CS76640	Oceanic climate	western_europe
<input checked="" type="checkbox"/>	9506	IP-Alb-0	CS76662	Mediterranean hot summer climates	admixed
<input checked="" type="checkbox"/>	9507	IP-Coa-0	CS76775	Mediterranean hot summer climates	spain
<input checked="" type="checkbox"/>	9508	IP-Mos-1	CS77108	Mediterranean hot summer climates	admixed
<input checked="" type="checkbox"/>	9509	IP-Reg-0	CS77207	Mediterranean hot summer climates	spain
<input checked="" type="checkbox"/>	9510	IP-Rei-0	CS77208	Mediterranean hot summer climates	spain
<input checked="" type="checkbox"/>	9511	IP-Vav-0	CS78835	Mediterranean hot summer climates	spain
<input checked="" type="checkbox"/>	9512	IP-Vid-1	CS78842	Mediterranean hot summer climates	spain
<input checked="" type="checkbox"/>	9513	IP-Adc-5	CS76646	Mediterranean hot summer climates	admixed

Compute results >

If you hover over any  symbol, additional information (e.g., for tabs or table headers) are shown.

[About](#)
[Select transcript and accessions](#)
[Results for selection](#)
[Search SNP database](#)

Results: AT4G30120.1 Download options

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)

[Protein haplotype <-> Accession table](#)

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.


Introns
 UTR
 Protein domains
 SNPs
 Syn. SNPs
 Nonsyn. SNPs
 CDS only
 Protein sequence only

Position: 0 Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

The figure is a genomic feature viewer for transcript AT4G30120.1. It displays a horizontal track with various genomic features. From top to bottom, the tracks are: Sequence (grey), 3'-UTR (red), 5'-UTR (red), CDS (blue), Introns (black), Nonsyn. SNPs (orange), NonCDS SNPs (green), Syn. SNPs (purple), and Prot. domains (grey). The x-axis represents genomic position from 0 to 3000. A red arrow points to the 'Check selected accessions' button. A green button labeled 'Download view' is located at the bottom right of the viewer.

2.4 Results for selection

If you compute the results SNPstar shifts directly to the section 'Results for selection'.



[About](#)
[Select transcript and accessions](#)
[Results for selection](#)
[Search SNP database](#)

Results: AT4G30120.1 Download options

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

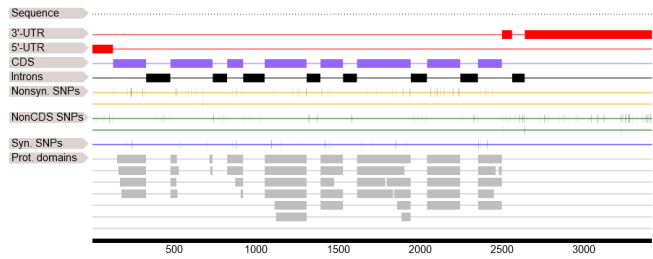
[Sequence viewer](#)
[SNPs](#)
[SNPs <-> Accessions](#)
[Haplotype <-> Accession table](#)

[Protein haplotype <-> Accession table](#)

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns
 UTR
 Protein domains
 SNPs
 Syn. SNPs
 Nonsyn. SNPs
 CDS only
 Protein sequence only

Position: 0 Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black



Here, you have different options to display your results as described in the following chapters.

2.4.1 Sequence viewer

The first tab of results is the 'Sequence viewer'.

i Top: Graphical overview of SNP positions in relation to nucleotide and amino acid sequence features like UTRs, intron, protein domains. Bottom: Plain genomic, coding and amino acid sequence with SNPs color coded by SNP type. Amino acid substitution (not context dependent but for single SNPs).

All options and possibilities are described in the following.

[About](#)
[Select transcript and accessions](#)
[Results for selection](#)
[Search SNP database](#)

Results: AT4G30120.1 Download options

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

➔
[Sequence viewer](#)
[SNPs](#)
[SNPs <-> Accessions](#)
[Haplotype <-> Accession table](#)

[Protein haplotype <-> Accession table](#)

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns
 UTR
 Protein domains
 SNPs
 Syn. SNPs
 Nonsyn. SNPs
 CDS only
 Protein sequence only

Position: 0 Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

You have different possible selections for 'Sequence viewer'. You can mark none or all of the available genomic features for display.

[About](#)
[Select transcript and accessions](#)
[Results for selection](#)
[Search SNP database](#)

Results: AT4G30120.1 Download options

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

➔
[Sequence viewer](#)
[SNPs](#)
[SNPs <-> Accessions](#)
[Haplotype <-> Accession table](#)

[Protein haplotype <-> Accession table](#)

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns
 UTR
 Protein domains
 SNPs
 Syn. SNPs
 Nonsyn. SNPs
 CDS only
 Protein sequence only

Position: 0 Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

nothing selected: shown is the whole sequence in base pairs (bp) and coding sequence (CDS)

Introns: shown is the whole sequence (in bp), CDS and introns of selected gene

UTR (untranslated region): shown is the whole sequence (in bp), 3'-UTR, 5'-UTR and coding sequence (CDS) of selected gene

Prot. domains (protein domains): shown is the whole sequence (in bp), coding sequence (CDS) and protein domains of selected gene

Noncoding SNPs (Noncoding single nucleotide polymorphisms): shown is the whole sequence (in bp), coding sequence (CDS) and SNPs in non coding regions (NonCDS SNPs)

Syn. SNPs (Synonymous SNPs): shown is the whole sequence (in bp), coding sequence (CDS) and synonymous SNPs in coding regions

Nonsyn. SNPs (Nonsynonymous SNPs): shown is the whole sequence (in bp), coding sequence (CDS) and nonsynonymous SNPs in coding regions

CDS only (coding sequence only): shown is the whole sequence (in bp), nonsynonymous SNPs, synonymous SNPs and protein domains

Protein sequence only: shown is the whole sequence (in amino acids [AA]), nonsynonymous SNPs, synonymous SNPs and protein domains

To show multiple features, mark all elements of interest. As an example shown are introns, protein domains and nonsynonymous SNPs.

Selected transcript and accessions Results for selection Search SNP database Build & Download GWAS matrices Manual About

Results: AT4G30120.1 All Download

ATHMA3; A. THALIANA HEAVY METAL ATPASE 3; HMA3; Heavy metal atpase 3

Check selected accessions

Sequence viewer SNPs SNPs <-> Accessions DNA Haplotypes <-> Accessions Protein haplotypes <-> Accessions

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns UTR Protein domains NSNPs Syn. SNPs Nonsyn. SNPs CDS only Protein sequence only

Position: 0 Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

Sequence
CDS
Introns
Nonsyn. SNPs
Prot. domains

500 1000 1500 2000 2500 3000

Download view

It's possible to zoom in. Click and hold the left mouse button to mark a region of interest.

[About](#)
[Select transcript and accessions](#)
[Results for selection](#)
[Search SNP database](#)

Results: AT4G30120.1 Download options ▼ [Download](#)

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) [SNPs](#) [SNPs <-> Accessions](#) [Haplotype <-> Accession table](#)

[Protein haplotype <-> Accession table](#)

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns UTR Protein domains SNPs Syn. SNPs Nonsyn. SNPs CDS only Protein sequence only

Position: 1931G Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

[Download view](#)

The zoom factor is shown as well.

[About](#)
[Select transcript and accessions](#)
[Results for selection](#)
[Search SNP database](#)

Results: AT4G30120.1 Download options ▼ [Download](#)

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) [SNPs](#) [SNPs <-> Accessions](#) [Haplotype <-> Accession table](#)

[Protein haplotype <-> Accession table](#)

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns UTR Protein domains SNPs Syn. SNPs Nonsyn. SNPs CDS only Protein sequence only

Position: 1868T Zoom: x 2.6 SNP with minor allele frequency <1% Grey, >1% Black

[Download view](#)

To zoom out (to zoom factor 1) click with the right mouse button on the figure.

SNPs are displayed as vertical lines. Gray color indicates a SNP with a minor allele frequency <1%, while and black vertical lines indicate SNPs with a minor allele frequency >1%.

About

Select transcript and accessions Results for selection Search SNP database

Results: AT4G30120.1 Download options Download

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Check selected accessions

Sequence viewer **i** SNPs **i** SNPs <-> Accessions **i** Haplotype <-> Accession table **i**

Protein haplotype <-> Accession table **i**

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns UTR Protein domains SNPs Syn. SNPs Nonsyn. SNPs CDS only Protein sequence only

Position: 1868T Zoom: x 2.6 SNP with minor allele frequency <1% Grey, >1% Black

Download view

If you hover with the mouse cursor over a SNP the sequence position (in bp) and specific base are shown (e.g. 1107T).

About

Select transcript and accessions Results for selection Search SNP database

Results: AT4G30120.1 Download options Download

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Check selected accessions

Sequence viewer **i** SNPs **i** SNPs <-> Accessions **i** Haplotype <-> Accession table **i**

Protein haplotype <-> Accession table **i**

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns UTR Protein domains SNPs Syn. SNPs Nonsyn. SNPs CDS only Protein sequence only

Position: 1107T Zoom: x 28.6 SNP with minor allele frequency <1% Grey, >1% Black

Download view

With the mouse cursor on a specific SNP the position in bp (1107 - 1108), chromosome:whole chromosome position (4:14732522), polymorphism with CDS position in bp (C981T) and minor allele frequency (0.18%) are shown.



You can download the 'Sequence viewer' figure (as svg file) by clicking the 'Download view' button.



Beneath of the above described figure, you can find DNA or amino acid sequence views.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)

Results: AT4G30120.1 Download options ▼ [Download](#)

ATHMA3; A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)

[Protein haplotype <-> Accession table](#)

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns | UTR | Protein domains | SNPs | Syn. SNPs | Nonsyn. SNPs | CDS only | Protein sequence only

Position: 424G Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

[Download view](#)

3415 **Sequence**

Sequence type: Genomic sequence

```

1  ATAGAGACTG AAGTTTCAA ATGTTATATA TAATAATGTA AGCCTAAACC AAATAATTAA TTTCCTCCG ACATCATC GTAAAAGTAT CTCCTCGAAA
181 TTTCTCCGA GATCTCAAGC TCGTCGATGG GGAAGGTGA AGAGTCAAG AGAGTGAATT TACAGACAAG TTACTTCGAC GTCGTTGGAA TCTGCTGTTCC
281 ATCAAGGTT TCTCGTAG GTACGTTCT CCGTCGTTA AACAAATTCTC AGTCAAGTT CCTCTAGAA CCGTCATCGT TGTCCACGAT
381 ACTTTTTTA TTTCTCCACT TCGATCGGT AAGTAACAGA AACGTGTATT TTTCATCGA CGTTTTGACT TCGCTCTGA ATCATTGGTC GATTGATGAG
  
```

There are three different sequence types available.

- genomic sequence:** shows the whole sequence region of selected gene
- CDS:** shows only coding sequence of selected gene
- Amino acid sequence:** shows AA sequence of selected protein

3415 **Sequence**

Sequence type: Genomic sequence

- Genomic sequence
- CDS
- Amino acid sequence

```

1  A AGCCTAAACC AAATAATTAA TTTCCTCCG ACATCATC GTAAAAGTAT CTCCTCGAAA
   A AGAGTCAAG AGAGTGAATT TACAGACAAG TTACTTCGAC GTCGTTGGAA TCTGCTGTTCC
   T CACGGGTTA AACAAATTCTC AGTCAAGTT CCTCTAGAA CCGTCATCGT TGTCCACGAT
   A AACGTGTATT TTTCATCGA CGTTTTGACT TCGCTCTGA ATCATTGGTC GATTGATGAG
481 GCAAAGGTTT ACGTTTCTT AAGTCAATT AAACCCCTT TTTTGTGTG TGTGTGATA ATTGATTCT TGAAGTCAA GGCCTCGAAT CAAGCAAGAC
581 TAGAAGCAAG TTAGACATA TGGAGAAA CAAGTTGAA GAGTCAATGG CCAAGCCCTT TCGCAATAG TTTCTGGTGA CTGGAGTTC TCTCCTCTCT
681 CAGTACTTT TATAGTCCGC TTGAATGGCT CATTATTGTT GCCGTGGTGG CTGGAGTTT CCCCACTCTT CTAAGAGCTG TTGCTTCAGT CACAAGGTTCC
781 CGCTGACGA TCAAGGCTCT CACTCTAATT GCTGGTAAAA ACAAACTCT TTTCCTCGAG CATCAGCTTA TGTTTTTTT GATGAGGTTA ATGAATCTGA
881 TTTGGTTAAT TTTTCAAC AGTGATAGCA ACGCTATGTA TGCAGGATT CAGGAAGCT GCTACAATTG TTTTTTTT CAGTTGGA GATTGGCTTG
981 AGTCTAGTGC TCTCATAAG GTTACTTAAT TAGGTTTTC TGTTCTTGA TGATTTTAA CAGTCATAGT CATACAAAAA TGTGAAAATC AGTTCTGGTT
1081 TTCAATAATT CTATCAGTTT ATCATTGTTT TGAAGTAACA ATTGTTTTC GGCAGCATA GTAATGTCAT CACTGATGAG CTTAGCCA CAGGACAG
1181 TGATCTGGG TACTGGACTA GAAGTTGATG TTGATGAGGT TGGGTAAC ACCGTTGTTT CAGTTAAAGC TTGGAGAAGT ATACCGATCG ATGGAGTTGT
1281 GGTGGATGGA AGCTGTGAT TGGATGAGAA AACATTTGACA GGGGAATCC TCCCTGTCTT AAACAGAGGA GTCAGACTG TTATGGCTGC AACCATAAAT
1381 CTTAAAGTGA CAAAACATA TCACTCTCTT GCTCTCTTAA AGTATGATCC TAAAGTTTTT AGCTCATTGC CAAGATTTTG TGTATGCTAA GGTATATGAA
1481 GGTGAAAGC TACAGBCTCT TCCCGGGACT GCGTGTGG AAAATGACT AAGCTTAG AAGAAAGTCA AAAAAGCCAA ACCAAAATCT AAAGGTTTAT
  
```

● Intron | ● CDS | ● 5'-UTR | ● 3'-UTR | ● Nonsynonymous SNP | ● Synonymous SNP | ● SNP in UTR or Intron

The different colors indicate (for all three sequence types):

- black letters:** intron sequence
- purple letters:** CDS
- red letters:** 5' and 3' UTR sequence
- yellow background:** Nonsynonymous SNP
- blue background:** Synonymous SNP
- green background:** a SNP in a UTR or intron region

2.4.2 SNPs

The next tab of results is 'SNPs'. This part describes the characteristics of all detected SNPs in detail.

SNPs ⓘ: Information about each polymorphic position in the selected transcript computed for the selected accessions (GWAS, always for all accessions); e.g. base frequencies, protein description, amino acid frequencies, GWAS results etc.

All options and possibilities are described in the following.

Select transcript and accessions | Results for selection | Search SNP database

Build & Download GWAS matrices

Results: AT4G30120.1 Download options Download

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Check selected accessions

Sequence viewer ⓘ | **SNPs ⓘ** | SNPs <-> Accessions ⓘ | Haplotype <-> Accession table ⓘ

Protein haplotype <-> Accession table ⓘ

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns UTR Protein domains SNPs Syn. SNPs Nonsyn. SNPs CDS only Protein sequence only

Position: 637T Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

Sequence
3'-UTR
5'-UTR
CDS
Introns
Nonsyn. SNPs
NonCDS SNPs
Syn. SNPs
Prot. domains

500 1000 1500 2000 2500 3000

Download view

SNP_ID (Genomic position): ⓘ Chromosome number:genomic position e.g. 4:14730221. SNP identification number. Shown are chromosome:bp position (4:14730221) of each SNP of the selected gene. The empty field below the header can be used for a specific SNP ID search.

Base frequencies: ⓘ You can sort according to the highest base frequency for each SNP and you can input a cutoff for the lowest allowed minor allele frequency. Shown are polymorphisms (base → base) and the frequencies (number behind) specific SNPs. The polymorphism with the same bases (e.g. C → C) is for the reference genotype (Col-0).

Genic position: ⓘ Position 1, is the A in the ATG start codon. The 5'-UTR has negative indices. For isoforms on the reverse strand we count the same position, but in the reverse complement. SNP base pair position of selected gene gDNA (genomic DNA).

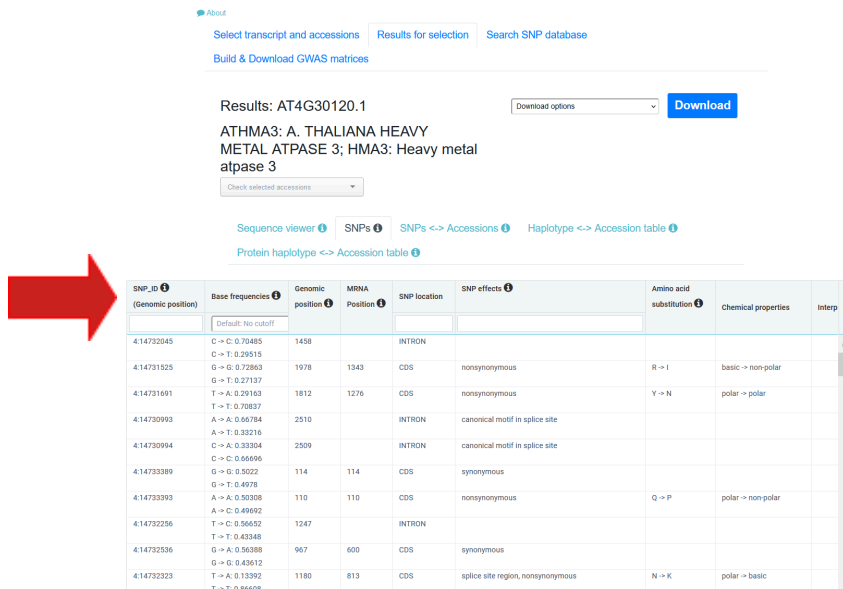
mRNA Position: ⓘ Position 1 is the A in the ATG start codon. The 5'-UTR has negative indices. Introns are not included. For isoforms on the reverse strand we count the same position, but in the reverse complement. SNP base pair position of selected gene mRNA.

SNP location: Gene structure location (e.g. 3' UTR (untranslated region), INTRON etc.) of SNP and selected gene.

SNP effects: ⓘ This includes the effect of the SNP at this position in the reference, ignoring other SNPs in the neighborhood. Canonical motif in splice site means the first dinucleotides in the intron of the splice site (donor GT, acceptor AG). Splice site region describes the region 1 – 3 in the flanking exon and the positions 3 – 8 in the flanking intron of the splice site. Effects are equally defined to SnpEff (A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3., Cigolani P; Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu W, Ruden DM. Fly (Austin): 2012). Effect of SNP on AA sequence of selected gene (e.g. stop codon (premature stop codon), nonsynonymous (AA substitution), canonical motif in splice site etc.)

Amino acid substitution: ⓘ Amino acid exchange caused by a nonsynonymous SNP in the codon (example: Q -> H means a glutamine in the reference sequence was substituted by a histidine in the SNP harboring allele).

Chemical properties: Physicochemical properties of AA substitution (e.g. non-polar → polar, basic → acidic etc.)



Results: AT4G30120.1 Download options Download

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Sequence viewer ⓘ SNPs ⓘ SNPs <-> Accessions ⓘ Haplotype <-> Accession table ⓘ Protein haplotype <-> Accession table ⓘ

SNP_ID ⓘ (Genomic position)	Base frequencies ⓘ	Genomic position ⓘ	MRNA Position ⓘ	SNP location	SNP effects ⓘ	Amino acid substitution ⓘ	Chemical properties	Interp
414732045	C → C: 0.70485 C → T: 0.29515	1458		INTRON				
414731525	G → G: 0.72863 G → T: 0.27137	1978	1343	CDS	nonsynonymous	R → I	basic → non-polar	
414731691	T → A: 0.29163 T → T: 0.70837	1812	1276	CDS	nonsynonymous	Y → N	polar → polar	
414730993	A → A: 0.66784 A → T: 0.33216	2510		INTRON	canonical motif in splice site			
414730994	C → A: 0.33304 C → C: 0.66696	2509		INTRON	canonical motif in splice site			
414733389	G → G: 0.5022 G → T: 0.4978	114	114	CDS	synonymous			
414733393	A → A: 0.50308 A → C: 0.49692	110	110	CDS	nonsynonymous	Q → P	polar → non-polar	
414732256	T → C: 0.56532 T → T: 0.43468	1247		INTRON				
414732536	G → A: 0.56388 G → G: 0.43612	967	600	CDS	synonymous			
414732323	T → A: 0.13392 T → T: 0.86608	1180	813	CDS	splice site region, nonsynonymous	N → K	polar → basic	

Interpro ID: Protein domain ID in which the AA substitutions are located. This information is based on the InterPro database (functional analysis of proteins by classifying them into families and predicting domains and important sites).

Interpro description: Description of InterPro domain

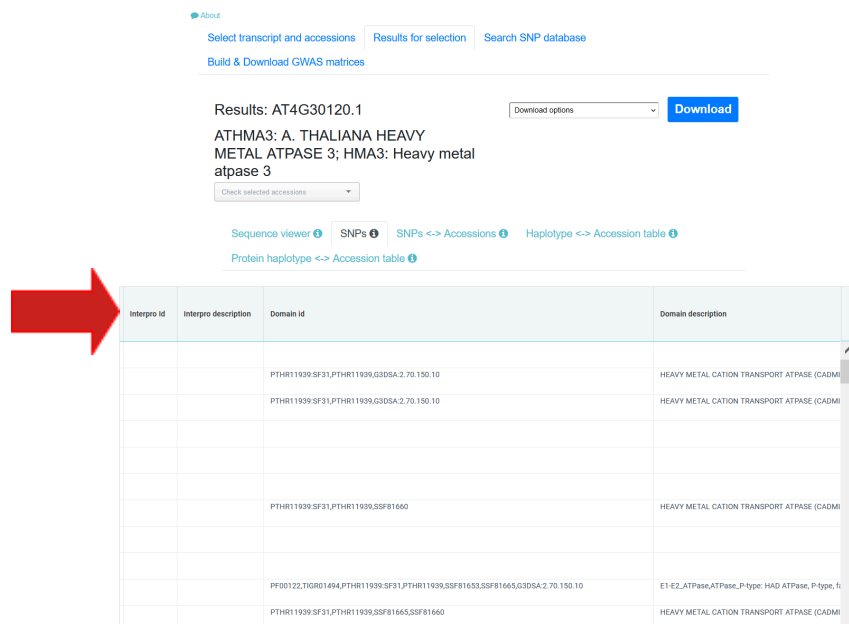
Domain id: Protein domain ID in which the AA substitutions are located. This information is based on Interproscan.

Domain description: Description of domain.

[About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#)
[Build & Download GWAS matrices](#)


Results: AT4G30120.1 Download options
 ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) [SNPs](#) [SNPs <-> Accessions](#) [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)



Interpro id	Interpro description	Domain id	Domain description
		PTHR11939:SF31,PTHR11939:G3D5A.2.70.150.10	HEAVY METAL CATION TRANSPORT ATPASE (CADMI
		PTHR11939:SF31,PTHR11939:G3D5A.2.70.150.10	HEAVY METAL CATION TRANSPORT ATPASE (CADMI
		PTHR11939:SF31,PTHR11939:SSF81660	HEAVY METAL CATION TRANSPORT ATPASE (CADMI
		PF00122,TIGR01494,PTHR11939:SF31,PTHR11939:SSF81653,SSF81665,G3D5A.2.70.150.10	E1-E2_ATPase_ATPase_P-type: HAD ATPase, P-type, f
		PTHR11939:SF31,PTHR11939:SSF81665,SSF81660	HEAVY METAL CATION TRANSPORT ATPASE (CADMI

Strand: Strand direction of sequence (e.g. - and +).

AraGWAS top hits:  Top associations imported from the AraGWAS tool as of 10/11/2024. The score is $-\log_{10}(\text{p-value})$ of the hit. The link links to the study on AraGWAS. The table can be sorted according to the lowest score in each cell. e.g.


study ID: 29.528 (AraGWAS SNP)


AraGWAS Catalog link of study information of selected SNP and gene.

AraGWAS score ($-\log_{10}(\text{p-value})$) of selected SNP and gene.

AraGWAS Catalog link of GWAS information of selected SNP and gene.

AraGWAS data based on the AraGWAS Catalog, which is a publicly and manually curated database for standardized GWAS (genome wide association study) results for *Arabidopsis thaliana*. Database is published on the 1001 genome homepage (<https://aragwas.1001genomes.org/#/>).

GWAS corrected p-values:  GWAS analysis was performed with the software GEMMA. Every climate factor was used as a phenotype and for every climate factor, one analysis run was carried out. Only SNPs with a minor allele frequency of 1% or higher were analysed. The obtained p-values were corrected for multiple comparisons using the Benjamini-Yekutieli method. In this column you can see all the correlations between SNPs and the 200 climate factors which have a corrected p-value lower than or equal to the threshold. You can sort the column according to the lowest p-value in each cell.

Number of associated climate factors:  Number of associated climate factors with a corrected P-value in the GWAS lower than or equal to the threshold.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 Download options

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Strand	AraGWAS top hits	GWAS corrected p-values	Number of associated climate factors
	P-value cutoff def.: <0.05	P-value cutoff def.: <0.05	
-	study_ID: 287: 29.528 (AraGWAS SNP)		0
-	study_ID: 287: 24.146 (AraGWAS SNP)		0
-	study_ID: 287: 21.205 (AraGWAS SNP)		0
-	study_ID: 287: 20.745 (AraGWAS SNP)		0
-	study_ID: 287: 20.745 (AraGWAS SNP)		0
-	study_ID: 287: 17.872 (AraGWAS SNP)		0
-	study_ID: 287: 17.872 (AraGWAS SNP)		0
-	study_ID: 287: 13.597 (AraGWAS SNP)		0
-	study_ID: 287: 11.792 (AraGWAS SNP)		0
-	study_ID: 287: 9.325 (AraGWAS SNP)		0

Calcium ATPase, transmembrane domain M.pro description

2.4.3 SNPs ↔ Accessions

The tab 'SNPs ↔ Accessions' shows a table for SNP/accession associations.

SNPs ← Accessions ⓘ : Information about each polymorphism in the selected transcript computed for the selected accession. List of accessions in which the polymorphism exists, effect of the polymorphism, amino acid change and SNP quality. All options and possibilities are described in the following.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 Download options

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3


[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns UTR Protein domains SNPs Syn. SNPs Nonsyn. SNPs CDS only Protein sequence only

Position: 0 Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

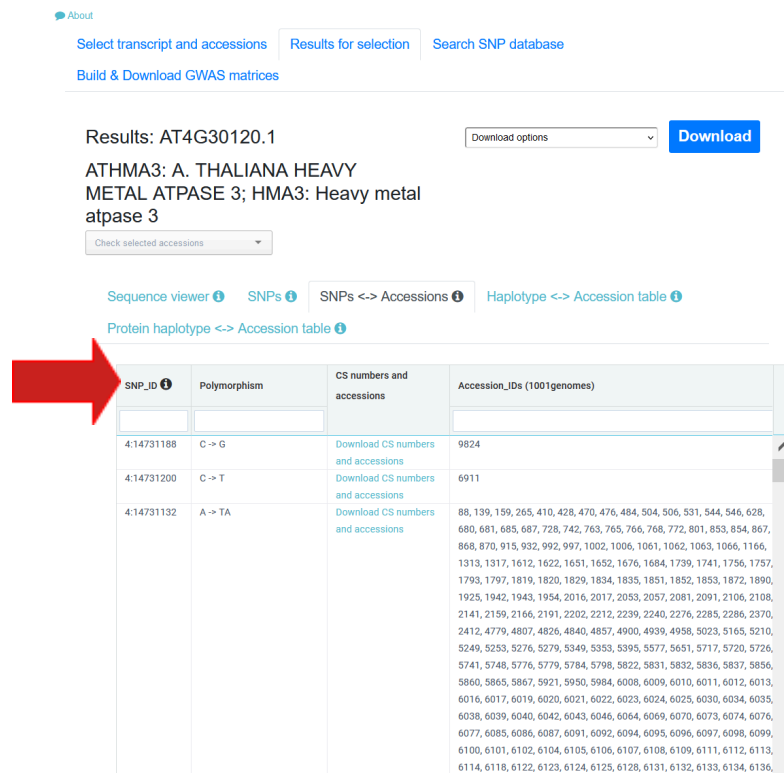
All empty fields beneath column titles can be used for specific searching (it is important to use the correct format!)

SNP Id:  Chr.Pos e.g. 4:14731188. SNP identification number. Shown are chromosome:bp positions (4:14731188) of each SNP of the selected gene.

Polymorphism: List of polymorphisms (SNPs or InDels). For example: C → G, always the first base (C) is from Col-0 (reference genotype) → the second base (G) from other accessions.

CS numbers and accessions: Download link for a list of genotypes different to Col-0. Shown are accessions with 1001 genome project Id, accession name and CS numbers.





Accession_IDs (1001 genomes): List of genotypes different to Col-0. Shown are accession IDs from 1001 Genome project.





Results: AT4G30120.1 Download options Download

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Check selected accessions


Sequence viewer  SNPs  SNPs <-> Accessions  Haplotype <-> Accession table 

Protein haplotype <-> Accession table 

SNP_ID 	Polymorphism	CS numbers and accessions	Accession_IDs (1001genomes)
4:14731188	C → G	Download CS numbers and accessions	9824
4:14731200	C → T	Download CS numbers and accessions	6911
4:14731132	A → TA	Download CS numbers and accessions	88, 139, 159, 265, 410, 428, 470, 476, 484, 504, 506, 531, 544, 546, 628, 680, 681, 685, 687, 728, 742, 763, 765, 766, 768, 772, 801, 853, 854, 867, 868, 870, 915, 932, 992, 997, 1002, 1006, 1061, 1062, 1063, 1066, 1166, 1313, 1317, 1612, 1622, 1651, 1652, 1676, 1684, 1739, 1741, 1756, 1757, 1793, 1797, 1819, 1820, 1829, 1834, 1835, 1851, 1852, 1853, 1872, 1890, 1925, 1942, 1943, 1954, 2016, 2017, 2053, 2057, 2081, 2091, 2106, 2108, 2141, 2159, 2166, 2191, 2202, 2212, 2239, 2240, 2276, 2285, 2286, 2370, 2412, 4779, 4807, 4826, 4840, 4857, 4900, 4939, 4958, 5023, 5165, 5210, 5249, 5253, 5276, 5279, 5349, 5353, 5395, 5577, 5651, 5717, 5720, 5726, 5741, 5748, 5776, 5779, 5784, 5798, 5822, 5831, 5832, 5836, 5837, 5856, 5860, 5865, 5867, 5921, 5950, 5984, 6008, 6009, 6010, 6011, 6012, 6013, 6016, 6017, 6019, 6020, 6021, 6022, 6023, 6024, 6025, 6030, 6034, 6035, 6038, 6039, 6040, 6042, 6043, 6046, 6064, 6069, 6070, 6073, 6074, 6076, 6077, 6085, 6086, 6087, 6091, 6092, 6094, 6095, 6096, 6097, 6098, 6099, 6100, 6101, 6102, 6104, 6105, 6106, 6107, 6108, 6109, 6111, 6112, 6113, 6114, 6118, 6122, 6123, 6124, 6125, 6128, 6131, 6132, 6133, 6134, 6136,

2.4.4 DNA Haplotype ↔ Accessions

The tab 'DNA Haplotype ↔ Accessions' shows a table for DNA haplotype/accession associations.

DNA Haplotype ↔ Accessions  : Accessions with the same polymorphisms in the selected CDS of the selected transcript have the same transcript specific haplotype. All options and possibilities are described in the following.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 Download options v Download

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns | UTR | Protein domains | SNPs | Syn. SNPs | Nonsyn. SNPs | CDS only | Protein sequence only

Position: 0 Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

Download view

All empty fields beneath column titles can be used for specific searching (use the correct format!)

Haplotype color: Color of each DNA haplotype which is shown in the map of all haplotypes (map is beneath the table!)

Haplotype ID: i Transcript Id.Haplotype number_dna (e.g. AT4G30120.1_dna).

AT4G30120.1 1 dna

Gene ID haplotype number type of sequence

Accessions: Count of accessions of each haplotype.

SNPs: Count of SNPs of each haplotype compared to reference (Col-0). The haplotype group with a SNP number of 0 is the reference haplotype group (accession haplotype group with Col-0).

Sequence: Download link of CDS haplotype sequence of each haplotype (FASTA file).

SNP list: i A list of the SNPs/InDels that differentiate this haplotype from the haplotype of the reference accession (Col-0). List of each SNP per each haplotype. Shown are SNP IDs in chromosome:bp position (4:14732536) of each SNP.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 Download options


ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Haplotype color	Haplotype_ID	#Accessions	#SNPs	Sequence	SNP list	CS ac
	AT4G30120.1.1_dna	88	0	Download Sequence		Dc
	AT4G30120.1.2_dna	1	2	Download Sequence	4:14732536,4:14733425	Dc
	AT4G30120.1.3_dna	1	5	Download Sequence	4:14732536,4:14733389,4:14733391,4:14733393,4:14733414	Dc
	AT4G30120.1.4_dna	1	5	Download Sequence	4:14732536,4:14733389,4:14733391,4:14733393,4:14733396	Dc
	AT4G30120.1.5_dna	1	5	Download Sequence	4:14731430,4:14732536,4:14733117,4:14733389,4:14733391	Dc
	AT4G30120.1.6_dna	1	4	Download Sequence	4:14732323,4:14732536,4:14733389,4:14733391	Dc
	AT4G30120.1.7_dna	4	4	Download Sequence	4:14731430,4:14732536,4:14733389,4:14733391	Dc

Map of all haplotypes

CS numbers and accessions: Download link for all accessions of specific haplotype. In the table 1001 genome ID, genotype name and CS number are shown.

Accession_IDs (1001 genomes):  The haplotype of one transcript might be the same for several accessions. So for each haplotype there is an accession list with the different accession IDs. The Accession_IDs originate from the 1001 genomes project. Genotype list of all haplotype groups of selected gene. Shown are accession IDs from the 1001 Genomes project.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 Download options

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

CS numbers and accessions	Accession_IDs (1001 genomes)
Download CS numbers and accessions	630,1070,5104,5486,5757,5768,5800,5811,6074,6088,6115,6137,6180,6191,6193,6909,6918,6919,6922,6945,6990,6997,7000,7062,7117,7125,7133,7161,7163,7165,7218,7231,7287,7288
Download CS numbers and accessions	403
Download CS numbers and accessions	7106
Download CS numbers and accessions	9554
Download CS numbers and accessions	10015
Download CS numbers and accessions	9954
Download CS numbers and accessions	6643,6968,9553,9847

Map of all haplotypes


Beneath of the 'DNA Haplotype ↔ Accessions' table a 'Map of all haplotypes' is shown.

[About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 Download options [Download](#)
 ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

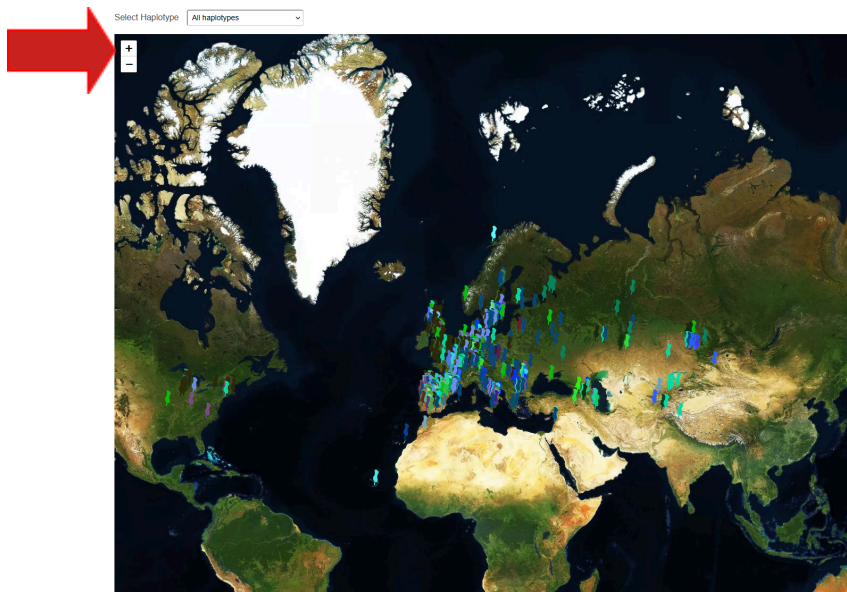
[Sequence viewer](#) [SNPs](#) [SNPs <-> Accessions](#) [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Haplotype color	Haplotype_ID	#Accessions	#SNPs	Sequence	SNP list	CS	ac
	AT4G30120.1.1_dna	88	0	Download Sequence		Dc	an
	AT4G30120.1.2_dna	1	2	Download Sequence	4:14732536,4:14733425	Dc	an
	AT4G30120.1.3_dna	1	5	Download Sequence	4:14732536,4:14733389,4:14733391,4:14733393,4:14733414	Dc	an
	AT4G30120.1.4_dna	1	5	Download Sequence	4:14732536,4:14733389,4:14733391,4:14733393,4:14733396	Dc	an
	AT4G30120.1.5_dna	1	5	Download Sequence	4:14731430,4:14732536,4:14733117,4:14733389,4:14733391	Dc	an
	AT4G30120.1.6_dna	1	4	Download Sequence	4:14732323,4:14732536,4:14733389,4:14733391	Dc	an
	AT4G30120.1.7_dna	4	4	Download Sequence	4:14731430,4:14732536,4:14733389,4:14733391	Dc	an

 Map of all haplotypes

The 'Map of all haplotypes' shows the geographic origin of all selected accessions. Each color is specific for a group of haplotypes. The zoom works via the buttons in the left upper corner or the mouse wheel. With holding the left mouse button and moving the mouse cursor the map can be shifted.

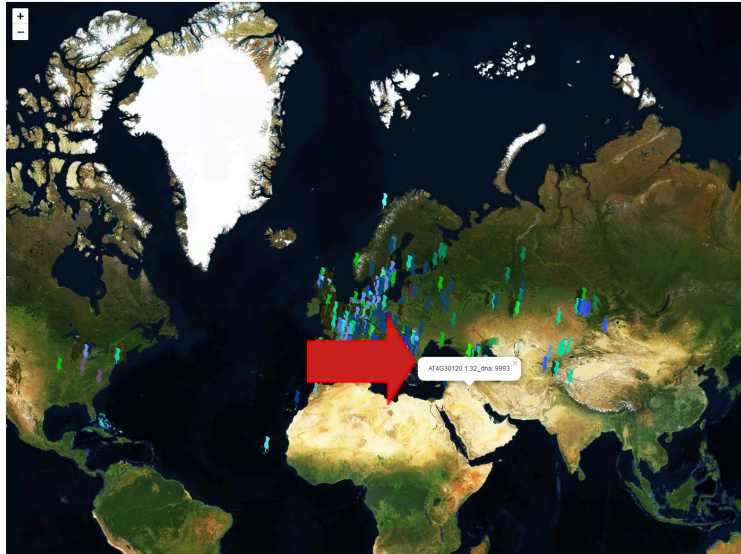
Map of all haplotypes



Left mouse click on a mark of the map shows haplotype and accession IDs (1001 genome ID). Both parameters are separated via a double colon (:).

Map of all haplotypes

Select Haplotype All haplotypes

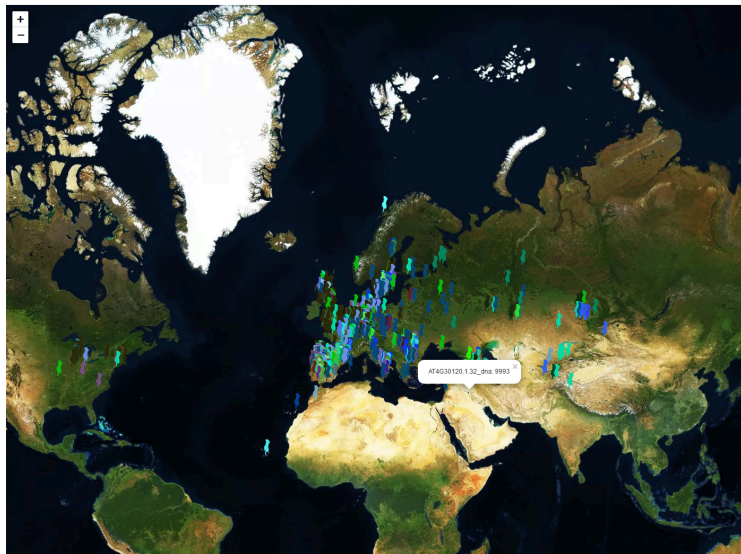


Select haplotype groups (accessions with the same haplotype) via the 'Select Haplotype' option.

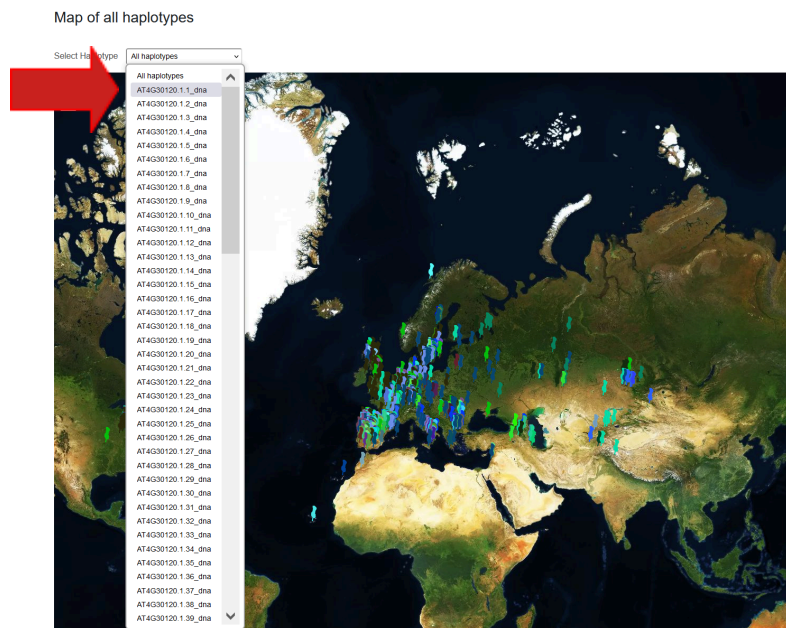


Map of all haplotypes

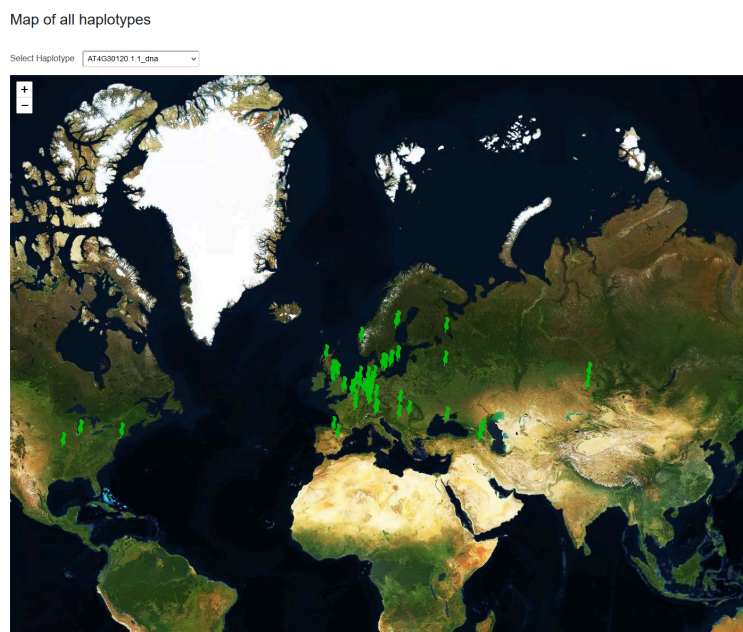
Select Haplotype All haplotypes



Selection of a specific haplotype (haplotype ID).



The geographic origins of all accessions within the haplotype AT4G30120.1.1_dna (haplotype ID) are shown as an example.



2.4.5 Proteotype ↔ Accession

The tab 'Proteotype ↔ Accessions' shows a table for protein haplotype/accession associations. The proteotype describes the protein haplotype.

Proteotype ← Accessions ⓘ : Accessions with polymorphisms in the selected CDS of the selected transcript that result in the same protein sequence have the same haplotype. All options and possibilities are described in the following.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 Download options v Download

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Sequence viewer ?
SNPs ?
SNPs <-> Accessions ?
Haplotype <-> Accession table ?

➔
Protein haplotype <-> Accession table ?

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns
 UTR
 Protein domains
 SNPs
 Syn. SNPs
 Nonsyn. SNPs
 CDS only
 Protein sequence only

Position: 1622T Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

Download view

All empty fields beneath column titles can be used for specific searching (use the correct format!)

Proteotype color: Color of each AA haplotype which is shown in the map of all haplotypes (map is beneath the table)

Proteotype ID: ? Transcript Id.Haplotype number_prot (e.g. AT4G30120.1.1_prot).
AT4G30120.1 1 prot

Gene ID haplotype number type of sequence

Accessions: Count of accessions of each haplotype.

SNPs: Count of SNPs of each proteotype compared to reference (Col-0). The proteotype group with a SNP number of 0 is the reference proteotype group (accession proteotype group with Col-0).

Sequence: Download link of AA sequence of each proteotype (FASTA file).

SNP list: ? A list of the SNPs/InDels that differentiate this proteotype from the proteotype of the reference accession (Col-0). List of SNPs causing AA substitutions per each proteotype. Shown are chromosome:bp positions (4:14732536) of each SNP.

[About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 Download options [Download](#)
 ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) [SNPs](#) [SNPs <-> Accessions](#) [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Haplotype color	Haplotype ID	#Accessions	#SNPs	Sequence	SNP list	CS acc
	AT4G30120.1_1_prot	172	0	Download Sequence		Dev anc
	AT4G30120.1_2_prot	1	5	Download Sequence	4.14732536,4.14733389,4.14733391,4.14733393,4.14733396	Dev anc
	AT4G30120.1_3_prot	1	7	Download Sequence	4.14731430,4.14731955,4.14732536,4.14733320,4.14733389,4.14733391,4.14733393	Dev anc
	AT4G30120.1_4_prot	11	6	Download Sequence	4.14731430,4.14732536,4.14733320,4.14733389,4.14733391,4.14733393	Dev anc
	AT4G30120.1_5_prot	4	6	Download Sequence	4.14731234,4.14732536,4.14733320,4.14733389,4.14733391,4.14733393	Dev anc
	AT4G30120.1_6_prot	3	5	Download Sequence	4.14732536,4.14733320,4.14733389,4.14733391,4.14733393	Dev anc
	AT4G30120.1_7_prot	1	6	Download Sequence	4.14731430,4.14732536,4.14733305,4.14733389,4.14733391,4.14733393	Dev anc

Map of all haplotypes

CS numbers and accessions: Download link for all accessions of specific haplotype. In the table 1001 genome Id, genotype name and CS number are shown.

Accession_IDs (1001 genomes): The proteotype of one transcript might be the same for several accessions. So far for each proteotype there is an accession list with the different accession IDs. The Accession_IDs originate from the 1001 genomes project. Genotype list of all proteotype groups of selected gene. Shown are accession IDs from the 1001 Genomes project.

[About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 Download options [Download](#)
 ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) [SNPs](#) [SNPs <-> Accessions](#) [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

CS numbers and accessions	Accession_IDs (1001 genomes)
Download CS numbers and accessions	403,410,424,630,801,1070,1158,1166,4807,5104,5486,5757,5768,5772,5800,5811,5837,5921,5984,5993,6074,6088,6094,6115,6137,6180,6191,6193,6680,6909,6918,6919,6922,6945,699
Download CS numbers and accessions	9554
Download CS numbers and accessions	9512
Download CS numbers and accessions	9507,9510,9511,9515,9522,9530,9541,9560,9577,9873,9943
Download CS numbers and accessions	9545,9600,9832,9871
Download CS numbers and accessions	9555,9743,9764
Download CS numbers and accessions	9710

Map of all haplotypes


Map of all proteotypes is shown beneath the table.

[About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#)
[Build & Download GWAS matrices](#)

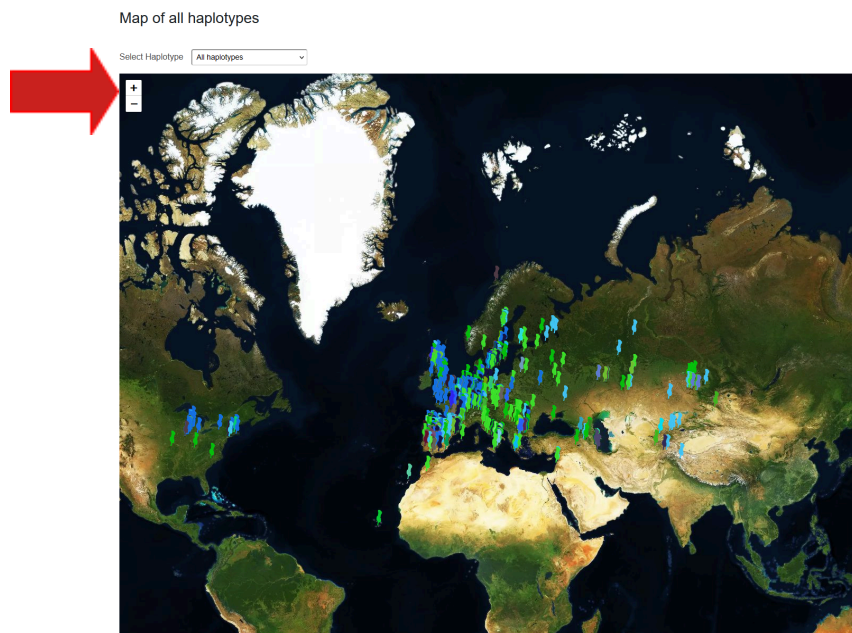
Results: AT4G30120.1 Download options [Download](#)
 ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) [SNPs](#) [SNPs <-> Accessions](#) [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Haplotype color	Haplotype ID	#Accessions	#SNPs	Sequence	SNP list	CS acc
	AT4G30120.1_1_prot	172	0	Download Sequence		Dev anc
	AT4G30120.1_2_prot	1	5	Download Sequence	4:14732536,4:14733389,4:14733391,4:14733393,4:14733396	Dev anc
	AT4G30120.1_3_prot	1	7	Download Sequence	4:14731430,4:14731955,4:14732536,4:14733320,4:14733389,4:14733391,4:14733393	Dev anc
	AT4G30120.1_4_prot	11	6	Download Sequence	4:14731430,4:14732536,4:14733320,4:14733389,4:14733391,4:14733393	Dev anc
	AT4G30120.1_5_prot	4	6	Download Sequence	4:14731234,4:14732536,4:14733320,4:14733389,4:14733391,4:14733393	Dev anc
	AT4G30120.1_6_prot	3	5	Download Sequence	4:14732536,4:14733320,4:14733389,4:14733391,4:14733393	Dev anc
	AT4G30120.1_7_prot	1	6	Download Sequence	4:14731430,4:14732536,4:14733305,4:14733389,4:14733391,4:14733393	Dev anc

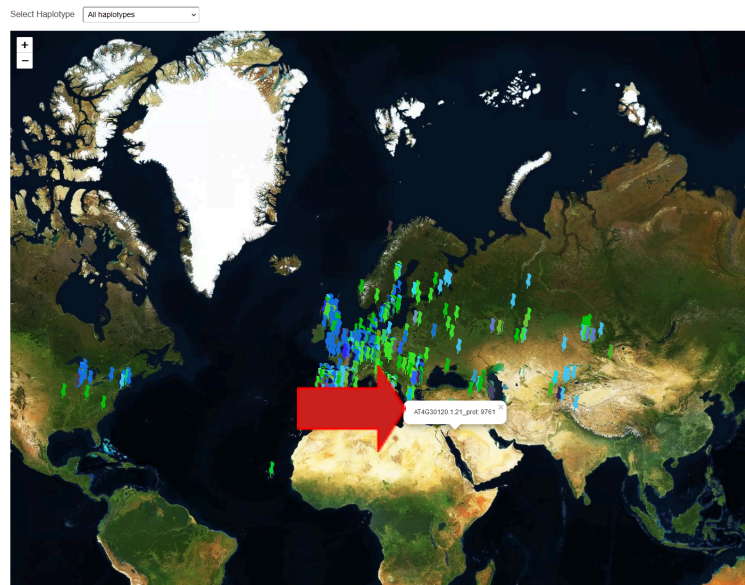
 Map of all haplotypes

The 'Map of all proteotypes' shows the geographical origin of all selected accessions. Each color is specific for a group of proteotypes. The zoom works via the buttons in the left upper corner or the mouse wheel. With holding the left mouse button and moving the mouse cursor the map can be shifted.

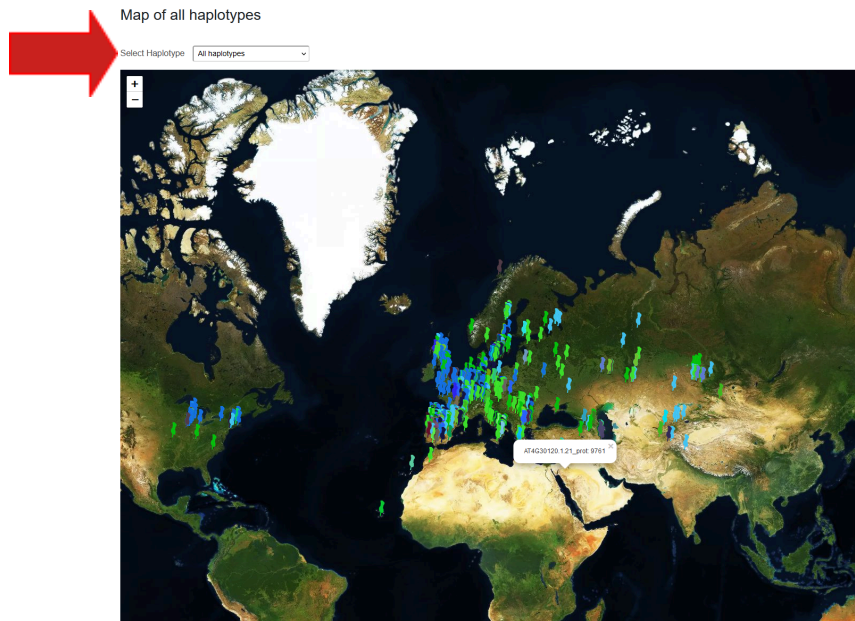


Left mouse click on a mark of the map shows protein haplotype and accession IDs (1001 Genomes ID). Both parameters are separated via a space.

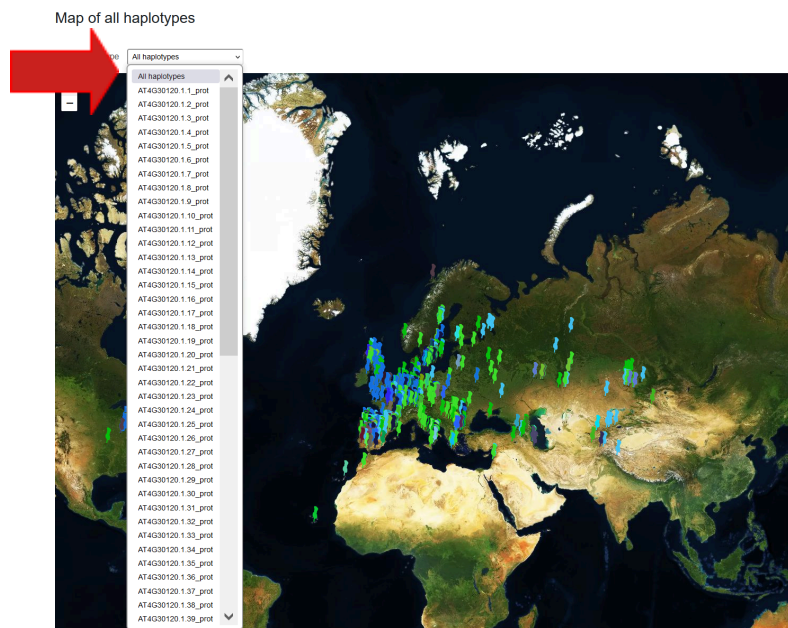
Map of all haplotypes



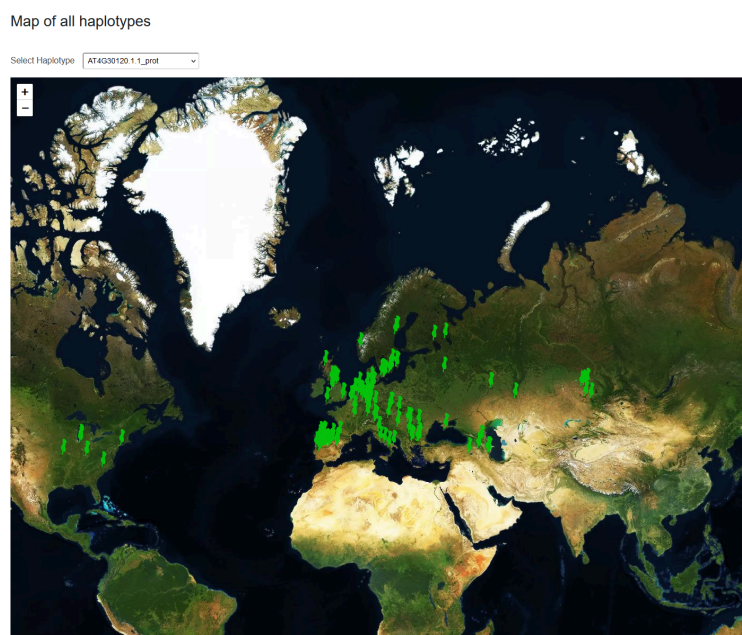
Select protein haplotype groups (accessions with the same protein haplotype) via 'Select Proteotype' option.



Selection of a specific proteotype (prototype ID).



The origin of all accessions from the protein haplotype AT4G30120.1_1_prot group (proteotype ID) are shown as an example.



2.4.6 Download options

In this chapter we describe download options for the data from the result tables.

2.4.6.1 Download all

To download all available result tables, open the drop-down menu on the upper right corner of the 'Results for selection' tab. Select the option 'All'.

About

Select transcript and accessions Results for selection Search SNP database

Results: AT4G30120.1

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Check selected accessions

Sequence viewer SNPs SNPs <-> Accessions Protein haplotype <-> Accession table

Download options Download

- All
- SNPs
- SNPs <-> Accessions
- DNA haplotype <-> Accessions table
- DNA haplotype fasta
- Protein haplotype <-> Accessions table
- Protein haplotype fasta
- Accession information

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns UTR Protein domains SNPs Syn. SNPs Nonsyn. SNPs CDS only Protein sequence only

Position: 0 Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

Download view

Press the 'Download' button.

About

Select transcript and accessions Results for selection Search SNP database

Results: AT4G30120.1

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Check selected accessions

Sequence viewer SNPs SNPs <-> Accessions Haplotype <-> Accession table

Protein haplotype <-> Accession table

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns UTR Protein domains SNPs Syn. SNPs Nonsyn. SNPs CDS only Protein sequence only

Position: 1522C Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

Download view

For all other download options select your file of interest and download it as described above.

For the description of the download options and the corresponding output files, see the following manual sections:

- For 'DNA haplotype ↔ Accession table' download see chapter 2.4.6.2.
- For 'Protein haplotype fasta' download see chapter 2.4.6.3.

- For 'DNA haplotype fasta' download see chapter 2.4.6.4.
- For 'Protein haplotype ↔ Accession table' download see chapter 2.4.6.5.
- For 'SNPs' download see chapter 2.4.6.6.
- For 'SNPs ↔ Accessions' download see chapter 2.4.6.7.
- For 'Accession information' download see chapter 2.4.6.8.

2.4.6.2 DNA haplotype table

The 'dna_haplotype_table' is dumped as a csv file.

AT4G30120.1_dna_haplotype_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	199 KB
AT4G30120.1_haplotype_sequences_completeAA	17/10/2024 16:27	FASTA DNA	41 KB
AT4G30120.1_haplotype_sequences_completeC...	17/10/2024 16:27	FASTA DNA	187 KB
AT4G30120.1_protein_haplotype_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	50 KB
AT4G30120.1_snp_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	42 KB
AT4G30120.1_snp_accession_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	252 KB

DNA haplotype table shows:

haplotype_ID: AT4G30120.1_1_dna

Gene ID **haplotype number** **type of sequence**

number_of_accessions: Count of accessions of each haplotype.

list_of_snps: List of each SNP per each haplotype. Shown are chromosome:bp positions (4:14732536) of each SNP.

#	A	B	C
1	#all	accessions	selected
2	haplotype_ID	number_of_accessions	list_of_snps
3	AT4G30120.1_1_dna		88
4	AT4G30120.1_2_dna		1,4:14732536,4:14733425
5	AT4G30120.1_3_dna		1,4:14732536,4:14733389,4:14733391,4:14733393,4:14733414
6	AT4G30120.1_4_dna		1,4:14732536,4:14733389,4:14733391,4:14733393,4:14733396
7	AT4G30120.1_5_dna		1,4:14731430,4:14732536,4:14733117,4:14733389,4:14733391
8	AT4G30120.1_6_dna		1,4:14732323,4:14732536,4:14733389,4:14733391
9	AT4G30120.1_7_dna		4,4:14731430,4:14732536,4:14733389,4:14733391
10	AT4G30120.1_8_dna		4,4:14732536,4:14733389,4:14733391
11	AT4G30120.1_9_dna		1,4:14732536,4:14733110,4:14733389,4:14733391
12	AT4G30120.1_10_dna		1,4:14731525,4:14731691,4:14732176,4:14733044,4:14733376

list_accession_IDs: All accessions belonging to the specified haplotype.
haplotype_sequence_complete_CDS: Coding sequence of haplotypes.
length_haplotype_sequence: Sequence length of haplotypes in bp.

	D	E	F
1			
2	list_accession_IDs	haplotype_sequence_completeCDS	length_haplotype_sequence
3	630,1070,5104,5486,5757,5768,5800,5811	ATGGCGGAAGGTGAAGAGTCAAAGA	1629
4	403	ATGGCGGAAGGTGAAGAGTCAAAGA	1629
5	7106	ATGGCGGAAGGTGAAGAGTCAAAGA	1629
6	9554	ATGGCGGAAGGTGAAGAGTCAAAGA	1629
7	10015	ATGGCGGAAGGTGAAGAGTCAAAGA	1629
8	9964	ATGGCGGAAGGTGAAGAGTCAAAGA	1629

2.4.6.3 Proteotype sequence complete AA

The 'proteotype_sequence_completeAA' is dumped as a FASTA file.

AT4G30120.1_dna_haplotype_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	199 KB
AT4G30120.1_haplotype_sequences_completeAA	17/10/2024 16:27	FASTA DNA	41 KB
AT4G30120.1_haplotype_sequences_completeC...	17/10/2024 16:27	FASTA DNA	187 KB
AT4G30120.1_protein_haplotype_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	50 KB
AT4G30120.1_snp_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	42 KB
AT4G30120.1_snp_accession_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	252 KB

This file contains all AA sequences of each selected haplotype.

```
>AT4G30120.1_1.prot
MAEGEESKKNLQTSYFDVVGICCSSEVSIQVNLRLDGVGKFEVIVPSRTVIVVHDTFFISPLQIVKALNQARLEASVRYGETLSKQWSPFPAIVSGVLLVLSFFKVFYSPLEWLAIVAVAG
VFPIAKAVASVTRFRFDINALLTAVIATCMQDFTEAATVFLFVADWLESAHAKASIVMSSLSLAPRKAVIADTGLEVDVDEINGINTVSVKAGESIPIDGVVDDGSDVDEKLTGESFPV
SKQRESTMATINLNGYKVKTTALARDCVAKMTKLVVEAQKSTKTQRFIDKCSRYTTPAVVSAACFVPIVLLKQDLSHWFHALVLSVGGPCGLLSTPVATFCALTAAKTSGLIKTG
DCLLETAKIKVAFDKTGTITKAEPVMSDFRSLSPSINLHLLYVSSIECKSSHPMAALIDYARSVSVPEKPDIVENFQNFPEGGVYGRIDGQDITGNKRIARAGCLTDNVPDIATMKRGT
IGIYMGAKLTGSFNLLDGCYVGAQALKEKS.
->AT4G30120.1_2.prot
MAEGEESKKNLQTSYFDVVGICCSSEVSIQVNLRLDGVGKFEVIVPSRTVIVVHDTFFISPLQIVKALNQARLEASVRYGETLSKQWSPFPAIVSGVLLVLSFFKVFYSPLEWLAIVAVAG
VFPIAKAVASVTRFRFDINALLTAVIATCMQDFTEAATVFLFVADWLESAHAKASIVMSSLSLAPRKAVIADTGLEVDVDEINGINTVSVKAGESIPIDGVVDDGSDVDEKLTGESFPV
SKQRESTMATINLNGYKVKTTALARDCVAKMTKLVVEAQKSTKTQRFIDKCSRYTTPAVVSAACFVPIVLLKQDLSHWFHALVLSVGGPCGLLSTPVATFCALTAAKTSGLIKTG
DCLLETAKIKVAFDKTGTITKAEPVMSDFRSLSPSINLHLLYVSSIECKSSHPMAALIDYARSVSVPEKPDIVENFQNFPEGGVYGRIDGQDITGNKRIARAGCLTDNVPDIATMKRGT
IGIYMGAKLTGSFNLLDGCYVGAQALKEKS.
->AT4G30120.1_3.prot
MAEGEESKKNLQTSYFDVVGICCSSEVSIQVNLRLDGVGKFEVIVPSRTVIVVHDTFFISPLQIVKALNQARLEASVRYGETLSKQWSPFPAIVSGVLLVLSFFKVFYSPLEWLAIVAVAG
VFPIAKAVASVTRFRFDINALLTAVIATCMQDFTEAATVFLFVADWLESAHAKASIVMSSLSLAPRKAVIADTGLEVDVDEINGINTVSVKAGESIPIDGVVDDGSDVDEKLTGESFPV
SKQRESTMATINLNGYKVKTTALARDCVAKMTKLVVEAQKSTKTQRFIDKCSRYTTPAVVSAACFVPIVLLKQDLSHWFHALVLSVGGPCGLLSTPVATFCALTAAKTSGLIKTG
DCLLETAKIKVAFDKTGTITKAEPVMSDFRSLSPSINLHLLYVSSIECKSSHPMAALIDYARSVSVPEKPDIVENFQNFPEGGVYGRIDGQDITGNKRIARAGCLTDNVPDIATMKRGT
IGIYMGAKLTGSFNLLDGCYVGAQALKEKS.
->AT4G30120.1_4.prot
MAEGEESKKNLQTSYFDVVGICCSSEVSIQVNLRLDGVGKFEVIVPSRTVIVVHDTFFISPLQIVKALNQARLEASVRYGETLSKQWSPFPAIVSGVLLVLSFFKVFYSPLEWLAIVAVAG
VFPIAKAVASVTRFRFDINALLTAVIATCMQDFTEAATVFLFVADWLESAHAKASIVMSSLSLAPRKAVIADTGLEVDVDEINGINTVSVKAGESIPIDGVVDDGSDVDEKLTGESFPV
SKQRESTMATINLNGYKVKTTALARDCVAKMTKLVVEAQKSTKTQRFIDKCSRYTTPAVVSAACFVPIVLLKQDLSHWFHALVLSVGGPCGLLSTPVATFCALTAAKTSGLIKTG
DCLLETAKIKVAFDKTGTITKAEPVMSDFRSLSPSINLHLLYVSSIECKSSHPMAALIDYARSVSVPEKPDIVENFQNFPEGGVYGRIDGQDITGNKRIARAGCLTDNVPDIATMKRGT
IGIYMGAKLTGSFNLLDGCYVGAQALKEKS.
->AT4G30120.1_5.prot
MAEGEESKKNLQTSYFDVVGICCSSEVSIQVNLRLDGVGKFEVIVPSRTVIVVHDTFFISPLQIVKALNQARLEASVRYGETLSKQWSPFPAIVSGVLLVLSFFKVFYSPLEWLAIVAVAG
VFPIAKAVASVTRFRFDINALLTAVIATCMQDFTEAATVFLFVADWLESAHAKASIVMSSLSLAPRKAVIADTGLEVDVDEINGINTVSVKAGESIPIDGVVDDGSDVDEKLTGESFPV
SKQRESTMATINLNGYKVKTTALARDCVAKMTKLVVEAQKSTKTQRFIDKCSRYTTPAVVSAACFVPIVLLKQDLSHWFHALVLSVGGPCGLLSTPVATFCALTAAKTSGLIKTG
DCLLETAKIKVAFDKTGTITKAEPVMSDFRSLSPSINLHLLYVSSIECKSSHPMAALIDYARSVSVPEKPDIVENFQNFPEGGVYGRIDGQDITGNKRIARAGCLTDNVPDIATMKRGT
IGIYMGAKLTGSFNLLDGCYVGAQALKEKS.
->AT4G30120.1_6.prot
MAEGEESKKNLQTSYFDVVGICCSSEVSIQVNLRLDGVGKFEVIVPSRTVIVVHDTFFISPLQIVKALNQARLEASVRYGETLSKQWSPFPAIVSGVLLVLSFFKVFYSPLEWLAIVAVAG
VFPIAKAVASVTRFRFDINALLTAVIATCMQDFTEAATVFLFVADWLESAHAKASIVMSSLSLAPRKAVIADTGLEVDVDEINGINTVSVKAGESIPIDGVVDDGSDVDEKLTGESFPV
SKQRESTMATINLNGYKVKTTALARDCVAKMTKLVVEAQKSTKTQRFIDKCSRYTTPAVVSAACFVPIVLLKQDLSHWFHALVLSVGGPCGLLSTPVATFCALTAAKTSGLIKTG
DCLLETAKIKVAFDKTGTITKAEPVMSDFRSLSPSINLHLLYVSSIECKSSHPMAALIDYARSVSVPEKPDIVENFQNFPEGGVYGRIDGQDITGNKRIARAGCLTDNVPDIATMKRGT
IGIYMGAKLTGSFNLLDGCYVGAQALKEKS.
```

2.4.6.4 Haplotype sequences complete CDS

The 'haplotype_sequence_completeCDS' is dumped as a FASTA file.

AT4G30120.1_dna_haplotype_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	199 KB
AT4G30120.1_haplotype_sequences_completeAA	17/10/2024 16:27	FASTA DNA	41 KB
AT4G30120.1_haplotype_sequences_completeC...	17/10/2024 16:27	FASTA DNA	187 KB
AT4G30120.1_protein_haplotype_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	50 KB
AT4G30120.1_snp_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	42 KB
AT4G30120.1_snp_accession_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	252 KB

This file contains all CDS sequences of each selected haplotype.

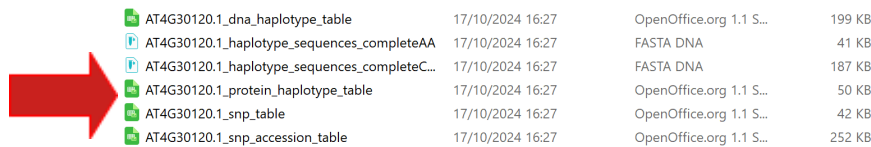
```

>AT4G30120.1.1_dna
ATGGCGAAGGTGAAGAGTCAAGAAGATGAATTTACAGACAAGTTACTTCGACGCTGTTGGAATCTGCTGTTTCATCGGAGGTTTCTATCTGAGGTAACGTTCTCCGTCAGGTGG;
TTCACTTAGCACTTGTAGTGTAGTAAGTGGTGTCCCTGTGGTCTTATCCTATCCACACCTGTTGCTACCTTTTGTGCTCTCACTAAGGCAGCCACGTCAGGGTTTCTGATCAA;
>AT4G30120.1.2_dna
ATGGCGAAGGTGAAGAGTCAAGAAGATGAATTTACAGACAAGTTACTTCGACGCTGTTGGAATCTGCTGTTTCATCGGAGGTTTCTATCTGAGGTAACGTTCTCCGTCAGGTGG;
TTCACTTAGCACTTGTAGTGTAGTAAGTGGTGTCCCTGTGGTCTTATCCTATCCACACCTGTTGCTACCTTTTGTGCTCTCACTAAGGCAGCCACGTCAGGGTTTCTGATCAA;
>AT4G30120.1.3_dna
ATGGCGAAGGTGAAGAGTCAAGAAGATGAATTTACAGACAAGTTACTTCGACGCTGTTGGAATCTGCTGTTTCATCGGAGGTTTCTACCGTAGGTAACGTTCTCCGTCAGGTGG;
TTCACTTAGCACTTGTAGTGTAGTAAGTGGTGTCCCTGTGGTCTTATCCTATCCACACCTGTTGCTACCTTTTGTGCTCTCACTAAGGCAGCCACGTCAGGGTTTCTGATCAA;
>AT4G30120.1.4_dna
ATGGCGAAGGTGAAGAGTCAAGAAGATGAATTTACAGACAAGTTACTTCGACGCTGTTGGAATCTGCTGTTTCATCGGAGGTTTCTATCTGAGGTAACGTTCTCCATCAACTTG;
TTCACTTAGCACTTGTAGTGTAGTAAGTGGTGTCCCTGTGGTCTTATCCTATCCACACCTGTTGCTACCTTTTGTGCTCTCACTAAGGCAGCCACGTCAGGGTTTCTGATCAA;
>AT4G30120.1.5_dna
ATGGCGAAGGTGAAGAGTCAAGAAGATGAATTTACAGACAAGTTACTTCGACGCTGTTGGAATCTGCTGTTTCATCGGAGGTTTCTATCTGAGGTAACGTTCTCCGTCAGGTGG;
TTCACTTAGCACTTGTAGTGTAGTAAGTGGTGTCCCTGTGGTCTTATCCTATCCACACCTGTTGCTACCTTTTGTGCTCTCACTAAGGCAGCCACGTCAGGGTTTCTGATCAA;
>AT4G30120.1.6_dna
ATGGCGAAGGTGAAGAGTCAAGAAGATGAATTTACAGACAAGTTACTTCGACGCTGTTGGAATCTGCTGTTTCATCGGAGGTTTCTATCTGAGGTAACGTTCTCCGTCAGGTGG;
TTCACTTAGCACTTGTAGTGTAGTAAGTGGTGTCCCTGTGGTCTTATCCTATCCACACCTGTTGCTACCTTTTGTGCTCTCACTAAGGCAGCCACGTCAGGGTTTCTGATCAA;
>AT4G30120.1.7_dna
ATGGCGAAGGTGAAGAGTCAAGAAGATGAATTTACAGACAAGTTACTTCGACGCTGTTGGAATCTGCTGTTTCATCGGAGGTTTCTATCTGAGGTAACGTTCTCCGTCAGGTGG;
TTCACTTAGCACTTGTAGTGTAGTAAGTGGTGTCCCTGTGGTCTTATCCTATCCACACCTGTTGCTACCTTTTGTGCTCTCACTAAGGCAGCCACGTCAGGGTTTCTGATCAA;

```

2.4.6.5 Proteotype table

The 'proteotype_table' is dumped as a csv file.



AT4G30120.1_dna_haplotype_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	199 KB
AT4G30120.1_haplotype_sequences_completeAA	17/10/2024 16:27	FASTA DNA	41 KB
AT4G30120.1_haplotype_sequences_completeC...	17/10/2024 16:27	FASTA DNA	187 KB
AT4G30120.1_protein_haplotype_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	50 KB
AT4G30120.1_snp_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	42 KB
AT4G30120.1_snp_accession_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	252 KB

The proteotype table shows:

proteotype_ID: **AT4G30120.1** **1** **prot**

Gene ID **proteotype number** **type of sequence**

number_of_accessions: count of accessions of each haplotype

list_of_snps: list of each SNP per each proteotype per gene. Shown are chromosome:bp positions (4:14732536) of each SNP

	A	B	C
1	#all	accessions	selected
2	haplotype_ID	number_of_accessions	list_of_snps
3	AT4G30120.1.1_prot	172	
4	AT4G30120.1.2_prot	1	4:14732536,4:14733389,4:14733391,4:14733393,4:14733396
5	AT4G30120.1.3_prot	1	4:14731430,4:14731955,4:14732536,4:14733320,4:14733389,4:14733391,4:14733393
6	AT4G30120.1.4_prot	11	4:14731430,4:14732536,4:14733320,4:14733389,4:14733391,4:14733393
7	AT4G30120.1.5_prot	4	4:14731234,4:14732536,4:14733320,4:14733389,4:14733391,4:14733393
8	AT4G30120.1.6_prot	3	4:14732536,4:14733320,4:14733389,4:14733391,4:14733393

list_accession_IDs: all accession belonging to the specified haplotype







proteotype_sequence_complete_AA: AA sequence of proteotype

length_proteotype_sequence: AA sequence length of proteotype

	D	E	F
1			
2	list_accession_IDs	haplotype_sequence_completeAA	length_haplotype_sequence
3	403,410,424,630,801,1070,1158,1166,4807,5104,5486,5757,578	MAEGEESKMMNLQTSYFDVVGICCSSEVSIVGNVLRQVDGVK	543
4		954 MAEGEESKMMNLQTSYFDVVGICCSSEVSIVGNVLRPLDGVK	543
5		9512 MAEGEESKMMNLQTSYFDVVGICCSSEVSIVGNVLRPLDGVK	543
6	9507,9510,9511,9515,9522,9530,9541,9560,9577,9873,9943	MAEGEESKMMNLQTSYFDVVGICCSSEVSIVGNVLRPLDGVK	543
7	9545,9600,9832,9871	MAEGEESKMMNLQTSYFDVVGICCSSEVSIVGNVLRPLDGVK	543
8	9555,9743,9764	MAEGEESKMMNLQTSYFDVVGICCSSEVSIVGNVLRPLDGVK	543

2.4.6.6 SNP table

The 'SNP_table' is dumped as a csv file.

	AT4G30120.1_dna_haplotype_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	199 KB
	AT4G30120.1_haplotype_sequences_completeAA	17/10/2024 16:27	FASTA DNA	41 KB
	AT4G30120.1_haplotype_sequences_completeC...	17/10/2024 16:27	FASTA DNA	187 KB
	AT4G30120.1_protein_haplotype_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	50 KB
	AT4G30120.1_snp_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	42 KB
	AT4G30120.1_snp_accession_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	252 KB

The SNP table shows:

SNP ID: SNP identification number. Shown are chromosome:bp position (4:14730221) of each SNP of selected gene.

Variant_0: Base pair variant on a specific position with the highest frequency of selected gene.

Frequency_0: Base frequency of variant 0 from selected gene.

Variant_1: Base pair variant on a specific position with the second highest frequency of selected gene.

Frequency_1: Base frequency of variant 1 from selected gene.

Variant_2: Base pair variant on a specific position with the third highest frequency of selected gene.

Frequency_2: Base frequency of variant 2 from selected gene.

SNP relative position mRNA: SNP base pair position of selected gene mRNA.

SNP relative position genome: SNP base pair position of selected gene gDNA (genomic DNA).

	A	B	C	D	E	F	G	H	I
1	#All accessions selected								
2	SNP id	Variant_0	Frequency_0	Variant_1	Frequency_1	Variant_2	Frequency_2	SNP relative position mRNA	SNP relative position genome
3	4:14730221	C -> C	0.99824	C -> A	0.00176	noval	noval	2463	3282
4	4:14730223	T -> T	0.98920	T -> C	0.01080	noval	noval	2461	3280
5	4:14730234	C -> C	0.99912	C -> A	0.00088	noval	noval	2450	3269
6	4:14730241	G -> T	0.70749	G -> G	0.29251	noval	noval	2443	3262
7	4:14730247	A -> A	0.99031	A -> T	0.00969	noval	noval	2437	3256
8	4:14730252	T -> G	0.67469	T -> T	0.32531	noval	noval	2432	3251
9	4:14730270	C -> C	0.99912	C -> T	0.00088	noval	noval	2414	3233
10	4:14730277	C -> C	0.99912	C -> A	0.00088	noval	noval	2407	3226
11	4:14730312	G -> G	0.99824	G -> A	0.00176	noval	noval	2372	3191
12	4:14730332	A -> A	0.99912	A -> G	0.00088	noval	noval	2352	3171
13	4:14730350	C -> C	0.9815	C -> T	0.0185	noval	noval	2334	3153
14	4:14730354	A -> A	0.95419	A -> G	0.04581	noval	noval	2330	3149
15	4:14730358	T -> T	0.99648	T -> A	0.00352	noval	noval	2326	3145
16	4:14730360	A -> A	0.99648	A -> T	0.00352	noval	noval	2324	3143

SNP location: Gene structure location (e.g. 3' UTR, INTRON etc.) of SNP and selected gene.

SNP effects: Effect of SNP on AA sequence of selected gene (e.g. stop codon (premature stop codon), nonsynonymous (AA substitution), canonical motif in splice site etc.)

Substitution_0: AA substitution on a specific position with the highest frequency of selected gene.

Substitution_frequency_0: AA substitution of substitution 0 from selected gene.

Substitution_1: AA substitution on a specific position with the second highest frequency of selected gene.

Substitution_frequency_1: AA substitution of substitution 1 from selected gene.

Substitution_properties: Physico-chemical properties of AA substitution (e.g. non-polar → polar, basic → acidic etc.)

Domain Id: Protein domain ID in which the AA substitutions are located.

	J	K	L	M	N	O	P	Q
1								
2	SNP location	SNP effects	Substitution_0	Substitution_frequency_0	Substitution_1	Substitution_frequency_1	Substitution_properties	Domain Id
51	INTRON	noval	noval	noval	noval	noval	noval	noval
52	INTRON	noval	noval	noval	noval	noval	noval	noval
53	INTRON	noval	noval	noval	noval	noval	noval	noval
54	3UTR	noval	noval	noval	noval	noval	noval	noval
55	3UTR	noval	noval	noval	noval	noval	noval	noval
56	3UTR	noval	noval	noval	noval	noval	noval	noval
57	3UTR	noval	noval	noval	noval	noval	noval	noval
58	3UTR	noval	noval	noval	noval	noval	noval	noval
59	3UTR	noval	noval	noval	noval	noval	noval	noval
60	CDS	stop codon	stop codon	frame shift	synonymous	noval	noval	noval
61	CDS	nonsynonymous	N -> K		0.00088	noval	noval	polar -> basic
62	CDS	synonymous	noval	noval	noval	noval	noval	noval
63	CDS	nonsynonymous	M -> T		0.00088	noval	noval	non-polar -> polar
64	CDS	synonymous	noval	noval	noval	noval	noval	noval
65	CDS	nonsynonymous	G -> D		0.00088	noval	noval	non-polar -> acidic
66	CDS	nonsynonymous	T -> N		0.00017	noval	noval	polar -> polar
67	CDS	synonymous	noval	noval	noval	noval	noval	noval
68	CDS	nonsynonymous	I -> T		0.00088	noval	noval	non-polar -> polar
69	CDS	splice site region	noval	noval	noval	noval	noval	noval
70	INTRON	canonical motif in splice site	noval	noval	noval	noval	noval	noval
71	INTRON	splice site region	noval	noval	noval	noval	noval	noval

Domain description: Description of domain.

	R
1	
2	Domain description
51	noval
52	noval
53	noval
54	noval
55	noval
56	noval
57	noval
58	noval
59	noval
60	noval
61	HEAVY METAL CATION TRANSPORT ATPASE (CADMIUM/ZINC TRANSPORTING ATPASE),CATION-TRANSPORTING ATPASE,HMA_2,HMA, heavy metal-associated domain
62	noval
63	HEAVY METAL CATION TRANSPORT ATPASE (CADMIUM/ZINC TRANSPORTING ATPASE),CATION-TRANSPORTING ATPASE,HMA_2,HMA, heavy metal-associated domain
64	noval
65	HEAVY METAL CATION TRANSPORT ATPASE (CADMIUM/ZINC TRANSPORTING ATPASE),CATION-TRANSPORTING ATPASE,HMA_2,HMA, heavy metal-associated domain
66	HEAVY METAL CATION TRANSPORT ATPASE (CADMIUM/ZINC TRANSPORTING ATPASE),CATION-TRANSPORTING ATPASE,HMA_2,HMA, heavy metal-associated domain
67	noval
68	HEAVY METAL CATION TRANSPORT ATPASE (CADMIUM/ZINC TRANSPORTING ATPASE),CATION-TRANSPORTING ATPASE,HMA_2,HMA, heavy metal-associated domain
69	noval
70	noval
71	noval

Strand: Strand direction (e.g. - and +) of sequence.

AraGWAS_score: AraGWAS score (-log₁₀ (p-value)) of selected SNP and gene.

AraGWAS_links: AraGWAS Catalog links of GWAS and study information of selected SNP and gene.

	S	T	U
1			
2	Strand	AraGWAS_score	AraGWAS_links
81	-	noval	noval
82	-	noval	noval
83	-	noval	noval
84	-	noval	noval
85	-	24.1462113508663	SNP_link: https://aragwas.1001genomes.org/#/study/287/associations/4_14731525 , study_link: https://aragwas.1001genomes.org/#/study/287 , score: 24.1
86	-	noval	noval
87	-	noval	noval
88	-	noval	noval
89	-	noval	noval
90	-	noval	noval
91	-	noval	noval
92	-	noval	noval
93	-	noval	noval
94	-	noval	noval
95	-	21.204987069462	SNP_link: https://aragwas.1001genomes.org/#/study/287/associations/4_14731691 , study_link: https://aragwas.1001genomes.org/#/study/287 , score: 21.2







GWAS corrected: GWAS analysis was performed with the software GEMMA. Every climate factor was used as a phenotype and for every climate factor, one analysis run was carried out. Only SNPs with a minor allele frequency of 1% or higher were analysed. The obtained p-values were corrected for multiple comparisons using the Benjamini-Yekutieli method. In this column you can see all the correlations between SNPs and the 200 climate factors which have a corrected p-value lower than or equal to the threshold. You can sort the column according to the lowest p-value in each cell.

Number of associated climate factors: Number of associated climate factors, with a corrected P-value in the GWAS lower than or equal to the threshold.

	V	W
1		
2	GWAS corrected	Number of associated climate factors
81	noval	0
82	noval	0
83	noval	0
84	noval	0
85	noval	0
86	noval	0
87	noval	0
88	noval	0
89	noval	0
90	noval	0
91	noval	0
92	noval	0
93	noval	0
94	noval	0
95	noval	0
96	noval	0

2.4.6.7 SNP accession table

The 'SNP_accession_table' is dumped as a csv file.

 AT4G30120.1_dna_haplotype_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	199 KB
 AT4G30120.1_haplotype_sequences_completeAA	17/10/2024 16:27	FASTA DNA	41 KB
 AT4G30120.1_haplotype_sequences_completeC...	17/10/2024 16:27	FASTA DNA	187 KB
 AT4G30120.1_protein_haplotype_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	50 KB
 AT4G30120.1_snp_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	42 KB
 AT4G30120.1_snp_accession_table	17/10/2024 16:27	OpenOffice.org 1.1 S...	252 KB

The SNP accession table shows:

SNP ID: List of SNP IDs of selected gene. Shown are chromosome:bp position (4:14731188) of each SNP

Polymorphism: List of polymorphisms (SNPs or InDels). For example: C → G, always the first base (C) is from Col-0 (reference genotype) and his group → the second base (G) from other accessions.

Genotypes: List of genotypes different to Col-0. Shown are accession IDs from 1001 Genome project.

CS number: List of genotypes different to Col-0. Shown are the CS accession IDs.

Accession: List of genotypes different to Col-0. Shown are accession names.

	A	B	C	D	E
1	#all accessions selected				
2	SNP id	Polymorphism	Genotypes	CS number	Accession
3	4:14731188	C -> G	9824	CS76702	Bes-5
4	4:14731200	C -> T	6911	CS76789	Cvi-0
5	4:14731132	A -> TA	88,139,159,265,410,428,	CS76790,CS77014,CS7	CYR,LDV-46,MAR2-3,PY
6	4:14731216	T -> C	6911	CS76789	Cvi-0
7	4:14731221	A -> T	9508,9518,9524,9528,95	CS77108,CS76660,CS7	IP-Mos-1,IP-Alm-0,IP-Ber
8	4:14731228	G -> A	9828	CS76721	Bra-0
9	4:14731234	C -> A	9533,9542,9545,9574,96	CS76763,CS76872,CS7	IP-Cem-0,IP-Fun-0,IP-He
10	4:14731260	T -> A	9543	CS76886	IP-Gra-0
11	4:14731261	T -> C	9879	CS77169	Per-0
12	4:14731275	T -> C	1158,1166,7169,8337,90	CS76656,CS77651,CS7	Aledal-6-49,Aledal-14-73,
13	4:14731393	G -> A	6090,7305,8343,9544,95	CS77288,CS78915,CS7	T1000,Pt-0,Na-1,IP-Gua-
14	4:14731395	GA -> A	9777	CS76880	Gn-1
15	4:14731396	G -> A	6897,9534,9841,9910	CS76430,CS76774,CS7	Ag-0,IP-Cmo-3,Ees-0,BR
16	4:14731430	T -> G	628,763,765,766,768,77	CS78951,CS76522,CS7	LI-OF-061,Kar-1,Sus-1,D
17	4:14731450	G -> A	9759,10013	CS76439,CS76388	Anz-0,Lerik1-3
18	4:14731471	AGAACTTTCC -> C	9099	CS76953	Istisu-9
19	4:14731480	A -> G	9991	CS76391	Vash-1

2.4.6.8 Accession information

The 'accession_information' is dumped as a csv file.

The accession information shows:

accession_id: Accession identification number based on the 1001 Genomes project.

accession: Name of accession.

cs_number: CS accession Ids.


country: Country of origin.

group: Population structure computed in 1001 genomes paper.

	A	B	C	D	E
1	accession_id	accession	cs_number	country	group
2	88	CYR	CS76790	FRA	western_europe
3	108	LDV-18	CS77013	FRA	western_europe
4	139	LDV-46	CS77014	FRA	western_europe
5	159	MAR2-3	CS77070	FRA	western_europe
6	265	PYL-6	CS77198	FRA	admixed
7	350	TOU-A1-88	CS77382	FRA	western_europe
8	351	TOU-A1-89	CS77383	FRA	western_europe
9	403	Zdarec3	CS78873	CZE	central_europe
10	410	Doubravnik7	CS76808	CZE	central_europe
11	424	Draha2	CS76812	CZE	central_europe
12	428	Borky1	CS76718	CZE	central_europe
13	430	Gr-1	CS76496	AUT	central_europe
14	470	BRR4	CS78943	USA	germany
15	476	BRR12	CS78944	USA	germany
16	484	BRR23	CS78945	USA	germany
17	504	BRR57	CS78946	USA	germany
18	506	BRR60	CS78947	USA	germany
19	531	BRR107	CS78948	USA	germany
20	544	LI-WP-039	CS78949	USA	germany
21	546	LI-WP-041	CS78950	USA	germany
22	628	LI-OF-061	CS78951	USA	western_europe
23	630	LI-OF-065	CS77036	USA	western_europe
24	680	LI-RR-096	CS78952	USA	germany
25	681	LI-RR-097	CS78953	USA	germany

2.5 Search SNP database

This database can be used to retrieve SNP information and accessions for SNPs. To get that information you need to identify SNP IDs of interest.



[About](#) |
 [Select transcript and accessions](#) |
 [Results for selection](#) |
 [Search SNP database](#) |
 [Build & Download GWAS matrices](#)

Results: AT4G30120.2 Accession information
ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

|
 [SNPs](#) |
 [SNPs <-> Accessions](#) |
 [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns |
 UTR |
 Protein domains |
 SNPs |
 Syn. SNPs |
 Nonsyn. SNPs |
 CDS only |
 Protein sequence only

Position: 0 **Zoom:** x 1 SNP with minor allele frequency <1% Grey, >1% Black



2.5.1 Get SNP information

[Manual](#) | [About](#)

[Select transcript and accessions](#) |
 [Results for selection](#) |
 [Search SNP database](#) |
 [Build & Download GWAS matrices](#) |
 SNP
PROT

|
 [Get accessions for SNPs](#)

1.5787
1.5808

Copy SNP ID or IDs of interest into the empty field.

About

Select transcript and accessions Results for selection Search SNP database

Build & Download GWAS matrices

Get SNP information Get accessions for SNPs

1:5787
1:5808
.
.
.

Download SNP information ⓘ

About

Select transcript and accessions Results for selection Search SNP database

Build & Download GWAS matrices

Get SNP information Get accessions for SNPs

4:14730221
4:14730223
4:14730234
4:14730241
4:14730247
|

Download SNP information ⓘ

Click 'Download SNP information'.

Download SNP information ⓘ : For each SNP ID in the list all the transcripts covering the SNP are listed. Additional information like SNP effect, SNP region and base frequency are added.

About

Select transcript and accessions Results for selection Search SNP database

Build & Download GWAS matrices

Get SNP information Get accessions for SNPs

4:14730221
4:14730223
4:14730234
4:14730241
4:14730247
|

Download SNP information ⓘ

Transcript ID: Shown are *A. thaliana* gene Ids and splicing variant (e.g. AT4G30120.1) in which the selected SNPs are located

SNP ID: SNP identification number. Shown are chromosome:bp positions (4:14730221) of each SNP of selected gene.

Variant_0: Base pair variant on a specific position with the highest frequency of selected gene.

Frequency_0: Base frequency of variant 0 from selected gene.

Variant_1: Base pair variant on a specific position with the second highest frequency of selected gene.

Frequency_1: Base frequency of variant 1 from selected gene.

SNP relative position mRNA: SNP base pair position of selected gene mRNA.

SNP relative position genome: SNP base pair position of selected gene gDNA (genomic DNA).

SNP location: Gene structure location (e.g. 3' UTR (untranslated region), INTRON etc.) of SNP and selected gene.

SNP effects: Effect of SNP on AA sequence of selected gene (e.g. stop codon (premature stop codon), nonsynonymous (AA substitution), canonical motif in splice site etc.)

	A	B	C	D	E	F	G	H	I	J
1	Transcript id	SNP id	Variant_0	Frequency_0	Variant_1	Frequency_1	SNP relative position mrna	SNP relative position genome	SNP location	SNP effects
2	AT4G30120.2	4:14730221	C -> C	0.99824	C -> A	0.00176	2296	2967	3UTR	noval
3	AT4G30120.2	4:14730223	T -> T	0.9859	T -> C	0.0141	2294	2965	3UTR	noval
4	AT4G30120.2	4:14730234	C -> C	0.99912	C -> A	0.00088	2283	2954	3UTR	noval
5	AT4G30120.2	4:14730241	G -> T	0.70749	G -> G	0.29251	2276	2947	3UTR	noval
6	AT4G30120.2	4:14730247	A -> A	0.99031	A -> T	0.00969	2270	2941	3UTR	noval
7	AT4G30120.1	4:14730221	C -> C	0.99824	C -> A	0.00176	2589	3408	3UTR	noval
8	AT4G30120.1	4:14730223	T -> T	0.9859	T -> C	0.0141	2587	3406	3UTR	noval
9	AT4G30120.1	4:14730234	C -> C	0.99912	C -> A	0.00088	2576	3395	3UTR	noval
10	AT4G30120.1	4:14730241	G -> T	0.70749	G -> G	0.29251	2569	3388	3UTR	noval
11	AT4G30120.1	4:14730247	A -> A	0.99031	A -> T	0.00969	2563	3382	3UTR	noval

Substitution_properties: Physico-chemical properties of AA substitution (e.g. non-polar → polar, basic → acidic etc.)

Domain ID: Protein domain ID in which the AA substitutions are located.

Domain description: Description of domain.

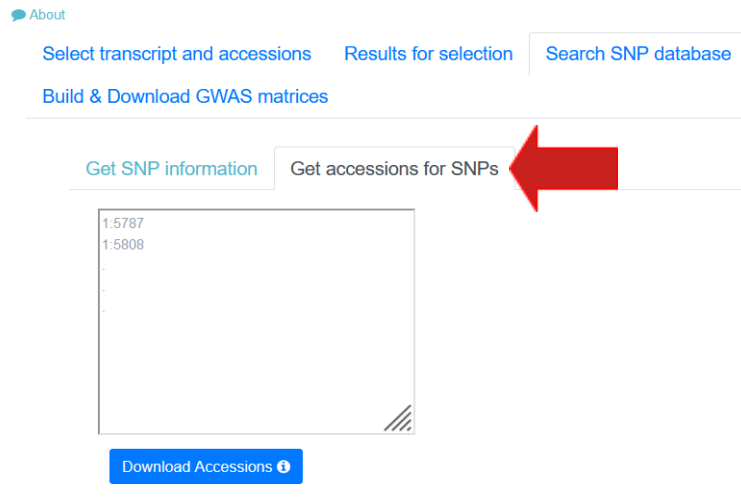
Strand: Strand direction (e.g. - and +) of sequence.

GWAS corrected: GWAS analysis was performed with the software GEMMA. Every climate factor was used as a phenotype and for every climate factor, one analysis run was carried out. Only SNPs with a minor allele frequency of 1% or higher were analysed. The obtained p-values were corrected for multiple comparisons using the Benjamini-Yekutieli method. In this column you can see all the correlations between SNPs and the 200 climate factors which have a corrected p-value lower than or equal to the threshold. You can sort the column according to the lowest p-value in each cell.

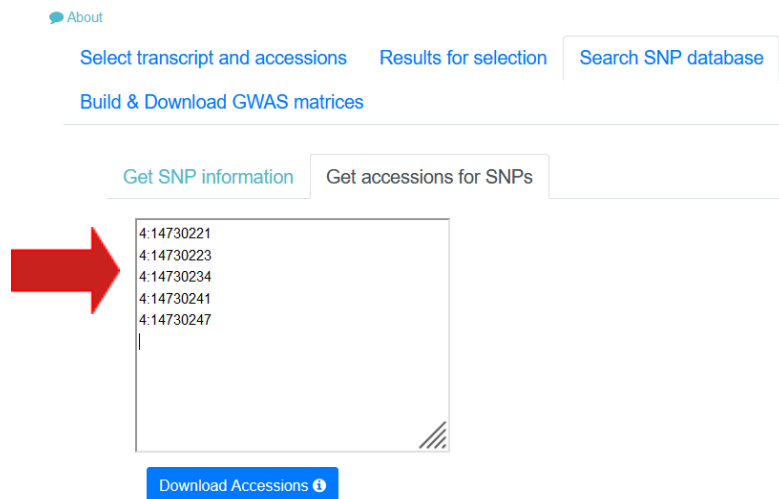
Number of associated climate factors: Number of associated climate factors, with a corrected P-value in the GWAS lower than or equal to the threshold.

	K	L	M	N	O	P	Q	R
1	Substitution properties	Domain id	Domain description	Strand	AraGWAS_score	AraGWAS_links	GWAS corrected	Number of associated climate factors
2	noval	noval	noval	-	noval	noval	noval	0
3	noval	noval	noval	-	noval	noval	noval	0
4	noval	noval	noval	-	noval	noval	noval	0
5	noval	noval	noval	-	noval	noval	noval	0
6	noval	noval	noval	-	noval	noval	noval	0
7	noval	noval	noval	-	noval	noval	noval	0
8	noval	noval	noval	-	noval	noval	noval	0
9	noval	noval	noval	-	noval	noval	noval	0
10	noval	noval	noval	-	noval	noval	noval	0
11	noval	noval	noval	-	noval	noval	noval	0

2.5.2 Get accessions for SNPs

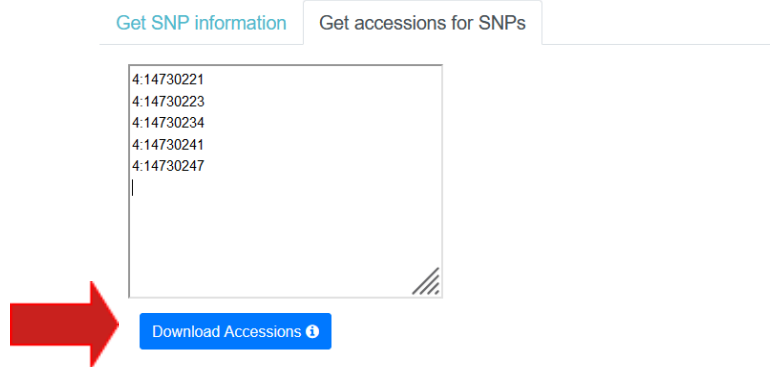


Copy and paste SNP ID or IDs of interest into the empty field.



Retrieve accessions for SNPs of interest by clicking 'Download Accessions'.

Download SNP information ⓘ : For each SNP ID in the list all the transcripts covering the SNP are listed. For each combination of transcript, SNP ID and variant all the genotypes with this specific variation are listed.



Transcript ID: Shown are *A. thaliana* gene IDs and splice variants (e.g. AT4G30120.2) in which the selected SNPs are located

SNP_ID: SNP identification number. Shown are chromosome:bp positions (4:14730221) of each SNP of selected gene.

T: Accessions with the base T on this position.

G: Accessions with the base G on this position.

A: Accessions with the base A on this position.

C: Accessions with the base C on this position.

	A	B	C	D	E	F
1	transcript_id	snp_id_c	T	G	A	C
2	AT4G30120.2:4:14730221	9608,9728,		88,108,139,159,265,350,351,403,410,424,428,430,470,476,484,504,506,531,544,546,628,630,68C		
3	AT4G30120.2:4:14730223			6074,8231,8235,8243,9646,9648,98,108,139,159,265,350,351,403,410,424,428,430,470,476,484,5		
4	AT4G30120.2:4:14730234	8230,		88,108,139,159,265,350,351,403,410,424,428,430,470,476,484,504,506,531,544,546,628,630,68C		
5	AT4G30120.2:4:14730241			265,350,351,403,424,428,430,470,476,484,504,506,531,544,546,628,630,68C		
6	AT4G30120.2:4:14730247	88,108,139,159,265,350,351,403,410,424,428,430,470,476,484,		*9102,9104,9106,9111,9113,9114,9115,9128,9130,9133,9134,		
7	AT4G30120.1:4:14730221	9608,9728,		88,108,139,159,265,350,351,403,410,424,428,430,470,476,484,504,506,531,544,546,628,630,68C		
8	AT4G30120.1:4:14730223			6074,8231,8235,8243,9646,9648,98,108,139,159,265,350,351,403,410,424,428,430,470,476,484,5		
9	AT4G30120.1:4:14730234	8230,		88,108,139,159,265,350,351,403,410,424,428,430,470,476,484,504,506,531,544,546,628,630,68C		
10	AT4G30120.1:4:14730241			265,350,351,403,424,428,430,470,476,484,504,506,531,544,546,628,630,68C		
11	AT4G30120.1:4:14730247	88,108,139,159,265,350,351,403,410,424,428,430,470,476,484,		*9102,9104,9106,9111,9113,9114,9115,9128,9130,9133,9134,		

2.6 Build & Download GWAS matrices

SNPstar enables the generation of GWAS matrices of interest. The computed matrix contains accessions and SNPs, DNA haplotypes or protein haplotypes.

[About](#)
[Select transcript and accessions](#)
[Results for selection](#)
[Search SNP database](#)

[Build & Download GWAS matrices](#)

Results: AT4G30120.2 Accession information [Download](#)

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) [SNPs](#) [SNPs <-> Accessions](#) [Haplotype <-> Accession table](#)

[Protein haplotype <-> Accession table](#)

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns UTR Protein domains SNPs Syn. SNPs Nonsyn. SNPs CDS only Protein sequence only

Position: 0 Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

The visualization shows a genomic track for the gene AT4G30120.2. The x-axis represents genomic position from 0 to 2500. The track includes:

- Sequence: A horizontal line representing the DNA sequence.
- 3'-UTR: A red bar at the end of the sequence.
- 5'-UTR: A red bar at the beginning of the sequence.
- CDS: A blue bar representing the coding sequence.
- Introns: Black bars representing non-coding regions within the CDS.
- Nonsyn. SNPs: Yellow bars representing non-synonymous single nucleotide polymorphisms.
- NonCDS SNPs: Green bars representing SNPs outside the coding sequence.
- Syn. SNPs: Grey bars representing synonymous SNPs.
- Prot. domains: Grey bars representing protein domains, with one labeled 'SSF81685'.

[Download view](#)

Select three different types of polymorphisms from the drop-down menu 'Select GWAS Matrix Type'.

[About](#)
[Select transcript and accessions](#)
[Results for selection](#)
[Search SNP database](#)

[Build & Download GWAS matrices](#)

Select GWAS Matrix Type:

Per default all transcripts are selected. Upload Transcript CSV File to select subset (IDs in each line): No file selected.

Select SNP Subset:

Select accessions:

No file selected.

Email for notification: [Compute GWAS Matrix](#)


'Select GWAS Matrix Type' options are 'SNP GWAS Matrix', 'Haplotype GWAS Matrix' and 'Proteotype GWAS Matrix'. If the 'SNP GWAS Matrix' is applied, a matrix for all available SNPs of selected genes is generated. For 'Haplotype GWAS Matrix' only DNA haplotypes (based on bases) and for 'Proteotype GWAS Matrix' only protein haplotypes (or proteotypes based on AAs).

[About](#)
[Select transcript and accessions](#)
[Results for selection](#)
[Search SNP database](#)

Build & Download GWAS matrices

Select GWAS Matrix Type:

- SNP GWAS Matrix
- Haplotype GWAS Matrix
- Proteotype GWAS Matrix

Per default  File to select subset (IDs in each line): No file selected.

Select SNP Subset:

Select accessions:

No file selected.

Email for notification:

You can upload a transcript table (csv file). If the default settings are used, a matrix for all transcripts (genes) is computed. If a download transcript file is used, a matrix for transcripts (genes) of interest is computed. The uploaded file needs to contain gene and transcript numbers (e.g. AT4G30120.1).


	A	B
1	AT4G30110.1	
2	AT4G30110.2	
3	AT4G30120.1	
4	AT4G30120.2	
5		

Click 'Browse' and select your transcript file of interest for upload.

[About](#)
[Select transcript and accessions](#)
[Results for selection](#)
[Search SNP database](#)

Build & Download GWAS matrices

Select GWAS Matrix Type:

Per default all transcripts are selected. Upload Transcript CSV File to select subs  No file selected.

Select SNP Subset:

Select accessions:

No file selected.

Email for notification:

To remove the selected transcript (Gene) file, click 'Unselect File/Default all transcripts'.

[About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#)
Build & Download GWAS matrices

Select GWAS Matrix Type:

Per default all transcripts are selected. Upload Transcript CSV File to select subset (IDs in each line) SNPstar_transcript.csv

Select SNP Subset:

Select accessions:
 No file selected.

Email for notification:

Next, select a SNP subset.

[About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#)
Build & Download GWAS matrices

Select GWAS Matrix Type:

Per default all transcripts are selected. Upload Transcript CSV File to select subset (IDs in each line) SNPstar_transcript.csv

Select SNP Subset:

Select accessions:
 No file selected.

Email for notification:

‘Select SNP Subsets’ options are ‘All SNPs’, ‘Coding region only’ and ‘Nonsynonymous SNPs only’. If ‘All SNPs’ are selected all SNPs are used. For ‘Coding region only’ SNPs in coding regions and for ‘Nonsynonymous SNPs only’ SNPs causing AA substitutions are considered.

[About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#)
Build & Download GWAS matrices

Select GWAS Matrix Type:

Per default all transcripts are selected. Upload Transcript CSV File to select subset (IDs in each line) SNPstar_transcript.csv

Select SNP Subset:

Email for notification:

If you would like to select accessions of interest you can upload an accession file. For the structure of that file see chapter 2.2.2. If the default settings are used a matrix for all accessions is computed. The upload works like for the transcript table (refer to 2.6). To remove the selected accession file use the ‘Unselect File/Default all accession’ button.

[About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#)
Build & Download GWAS matrices

Select GWAS Matrix Type:

Per default all transcripts are selected. Upload Transcript CSV File to select subset (IDs in each line): SNPstar_transcript.csv

Select SNP Subset:

Select accessions:
 SNPstar_accessions.csv

Email for notification:

Provide an email address for notification and press 'Compute GWAS Matrix'.

[About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#)
Build & Download GWAS matrices

Select GWAS Matrix Type:

Per default all transcripts are selected. Upload Transcript CSV File to select subset (IDs in each line): SNPstar_transcript.csv

Select SNP Subset:

Select accessions:
 SNPstar_accessions.csv

Email for notification:

If the calculation is in progress, an information window is displayed.

[Manual](#) [About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#) **Build & Download GWAS matrices**

Select GWAS Matrix Type:

Per default all transcripts are selected. Upload Transcript CSV File to select subset (IDs in each line): SNPstar_transcript.csv

Select SNP Subset:

Select accessions:
 SNPstar_accessions.csv

Email for notification:

After processing, download GWAS matrix at localhost:5000/query_result/gwas_matrix_snp_dbaf1dfec6478155759ea8a0078c35f0 (Currently Processing...)

To retrieve the GWAS matrix click on the green window (only when processing is finished).

[Manual](#) [About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#) **Build & Download GWAS matrices**

Select GWAS Matrix Type:

Per default all transcripts are selected. Upload Transcript CSV File to select subset (IDs in each line): SNPstar_transcript.csv

Select SNP Subset:


Select accessions:
 SNPstar_accessions.csv

Email for notification:

File ready: /query_result/gwas_matrix_snp_dbaf1dfec6478155759ea8a0078c35f0.zip


2.6.1 Layout of SNP GWAS Matrix

In the first horizontal row all selected accessions (IDs from the 1001 Genomes project) are shown.



	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	6145	6148	2053	6149	6150	6151	2057	6153	6154	6163	6166	8214	19950	6169	19951	6172
2	AT4G30110.1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
3	AT4G30110.2	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
4	AT4G30120.1	0	1	0	0	1	0	0	0	0	1	0	1	0	0	0
5	AT4G30120.2	0	1	0	0	1	0	0	0	0	1	0	1	0	0	0

In the first vertical column all selected transcripts are shown.




	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	6145	6148	2053	6149	6150	6151	2057	6153	6154	6163	6166	8214	19950	6169	19951	6172
2	AT4G30110.1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
3	AT4G30110.2	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
4	AT4G30120.1	0	1	0	0	1	0	0	0	0	1	0	1	0	0	0
5	AT4G30120.2	0	1	0	0	1	0	0	0	0	1	0	1	0	0	0

The matrix consists of numbers. The number zero (0) describes the status of Col-0. One (1) describes an allele different from Col-0. This matrix can be used for GWAS and/or kinship calculations.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	6145	6148	2053	6149	6150	6151	2057	6153	6154	6163	6166	8214	19950	6169	19951	6172	
2	AT4G30110.1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0
3	AT4G30110.2	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0
4	AT4G30120.1	0	1	0	0	1	0	0	0	0	0	1	0	1	0	0	0
5	AT4G30120.2	0	1	0	0	1	0	0	0	0	1	0	1	0	0	0	0


2.6.2 Layout of Haplotype GWAS Matrix

In the first horizontal row all selected accessions (IDs from the 1001 Genomes project) are shown.



	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	6145	6148	2053	6149	6150	6151	2057	6153	6154	6163	6166	8214	19950	6169	19951	6172
2	AT4G30110.1	68	68	33	73	33	70	33	33	33	33	26	70	10	70	9
3	AT4G30110.2	68	68	33	73	33	70	33	33	33	33	26	70	10	70	9
4	AT4G30120.1	38	38	36	38	36	47	36	36	36	36	84	84	87	84	80
5	AT4G30120.2	47	47	45	47	45	27	45	45	45	45	25	25	32	25	20

In the first vertical column all selected transcripts are shown.




	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	6145	6148	2053	6149	6150	6151	2057	6153	6154	6163	6166	8214	19950	6169	19951	6172
2	AT4G30110.1	68	68	33	73	33	70	33	33	33	33	26	70	10	70	9
3	AT4G30110.2	68	68	33	73	33	70	33	33	33	33	26	70	10	70	9
4	AT4G30120.1	38	38	36	38	36	47	36	36	36	36	84	84	87	84	80
5	AT4G30120.2	47	47	45	47	45	27	45	45	45	45	25	25	32	25	20

The matrix consists of numbers which describe a specific haplotype of specific accession and transcript. This matrix can be used for GWAS calculation.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		6145	6148	2053	6149	6150	6151	2057	6153	6154	6163	6166	8214	19950	6169	19951	6172
2	AT4G30110.1	68	68	33	73	33	70	33	33	33	33	33	26	70	10	70	9
3	AT4G30110.2	68	68	33	73	33	70	33	33	33	33	33	26	70	10	70	9
4	AT4G30120.1	38	38	36	38	36	47	36	36	36	36	36	84	84	87	84	80
5	AT4G30120.2	47	47	45	47	45	27	45	45	45	45	45	25	25	32	25	20


2.6.3 Layout of Proteotype GWAS Matrix

In the first horizontal row all selected accessions (IDs from the 1001 Genomes project) are shown.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		6145	6148	2053	6149	6150	6151	2057	6153	6154	6163	6166	8214	19950	6169	19951	6172
2	AT4G30110.1	36	36	30	41	30	36	30	30	30	30	30	30	36	68	36	70
3	AT4G30110.2	36	36	30	41	30	36	30	30	30	30	30	30	36	68	36	70
4	AT4G30120.1	57	57	54	57	54	67	54	54	54	54	54	21	21	19	21	19
5	AT4G30120.2	19	19	16	19	16	29	16	16	16	16	16	1	1	23	1	23

In the first vertical column all selected transcripts are shown.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		6145	6148	2053	6149	6150	6151	2057	6153	6154	6163	6166	8214	19950	6169	19951	6172
2	AT4G30110.1	36	36	30	41	30	36	30	30	30	30	30	30	36	68	36	70
3	AT4G30110.2	36	36	30	41	30	36	30	30	30	30	30	30	36	68	36	70
4	AT4G30120.1	57	57	54	57	54	67	54	54	54	54	54	21	21	19	21	19
5	AT4G30120.2	19	19	16	19	16	29	16	16	16	16	16	1	1	23	1	23

The matrix consists of numbers which describe a specific protein haplotype/proteotype of a specific accession and transcript. This matrix can be used for GWAS calculation.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		6145	6148	2053	6149	6150	6151	2057	6153	6154	6163	6166	8214	19950	6169	19951	6172
2	AT4G30110.1	36	36	30	41	30	36	30	30	30	30	30	30	36	68	36	70
3	AT4G30110.2	36	36	30	41	30	36	30	30	30	30	30	30	36	68	36	70
4	AT4G30120.1	57	57	54	57	54	67	54	54	54	54	54	21	21	19	21	19
5	AT4G30120.2	19	19	16	19	16	29	16	16	16	16	16	1	1	23	1	23

3 Use cases

SNPstar covers a number of different use cases. In this chapter, we will explain the most prominent use cases, which can be roughly grouped into single target gene, multiple target genes and GWAS.

3.1 Single target gene

In this first use case, we will consider a single target gene and all SNPs in this gene. We will again use *A. THALIANA HEAVY METAL ATPASE 3 (HMA3; AT4G30120)* as an example.

First, select gene and gene model (AT4G30120.1; see chapter 2.1.1). Then we select all available accessions (see chapter 2.2.1) and click 'Compute results' (see chapter 2.3) to get to the 'Sequence viewer' tab as displayed below.

[About](#)
[Select transcript and accessions](#) [Results for selection](#) [Search SNP database](#)

Results: AT4G30120.1 Download options
 ATHMA3: A. THALIANA HEAVY
 METAL ATPASE 3; HMA3: Heavy metal
 atpase 3

[Sequence viewer](#) [SNPs](#) [SNPs <-> Accessions](#) [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Please note: Indels are not included in this tab. For now the position in the feature viewer starts with 0, while the genomic position in the 5'UTR starts negative, with the A in the ATG start codon being position 1. The SNP labels in the feature viewer show the genomic position. For transcripts on the reverse strand the reverse complement is shown.

Introns UTR Protein domains SNPs Syn. SNPs Nonsyn. SNPs CDS only Protein sequence only

Position: 0 Zoom: x 1 SNP with minor allele frequency <1% Grey, >1% Black

3.1.1 Identification of SNPs in a single gene of interest

3.1.1.1 General selection and SNP identification

The Sequence viewer tab (see chapter 2.4.1) shows an overview of all SNPs represented in all accessions. Additional information are gene structure (e.g. CDS, Intron, etc.), detected protein domain locations and effect of SNPs (nonsynonymous, synonymous and non-CDS). Beneath the overview figure the sequence is shown.

For a closer look of detected SNPs use the 'SNPs' tab (see chapter 2.4.2). Different selection options can be used depending on your research interest. If you are interested in a specific SNP location use the column 'SNP location' for selection. As an example we use CDS for selection.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 All Download

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Check selected accessions](#) | [Sequence viewer](#) | [SNPs](#) | [Accessions](#) | [Haplotype <-> Accession table](#) | [Protein haplotype <-> Accession table](#)

SNP_ID (Genomic position)	Base frequencies	Genomic position	MRNA Position	SNP location	SNP effects	Amino acid substitution	Chemical properties	Interp
4.14732045	C -> C: 0.70485 C -> T: 0.29515	1458		INTRON				
4.14731525	G -> G: 0.72863 G -> T: 0.27137	1978	1343	CDS	nonsynonymous	R -> I	basic -> non-polar	
4.14731691	T -> A: 0.29163 T -> T: 0.70837	1812	1276	CDS	nonsynonymous	Y -> N	polar -> polar	
4.14730993	A -> A: 0.66784 A -> T: 0.33216	2510		INTRON	canonical motif in splice site			
4.14730994	C -> A: 0.33304 C -> C: 0.66696	2509		INTRON	canonical motif in splice site			
4.14733389	G -> G: 0.5022 G -> T: 0.4978	114	114	CDS	synonymous			
4.14733393	A -> A: 0.50308 A -> C: 0.49692	110	110	CDS	nonsynonymous	Q -> P	polar -> non-polar	
4.14732256	T -> C: 0.56652 T -> T: 0.43348	1247		INTRON				
4.14732536	G -> A: 0.50588 G -> G: 0.49412	967	600	CDS	synonymous			
4.14732323	T -> A: 0.13392 T -> T: 0.86608	1180	813	CDS	splice site region, nonsynonymous	N -> K	polar -> basic	

Type 'CDS' into the empty field below the column header. The selection works automatically (no need to press enter). Shown are all SNPs which are located in the CDS of the selected target gene.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 All Download

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Check selected accessions](#) | [Sequence viewer](#) | [SNPs](#) | [Accessions](#) | [Haplotype <-> Accession table](#) | [Protein haplotype <-> Accession table](#)

SNP_ID (Genomic position)	Base frequencies	Genomic position	MRNA Position	SNP location	SNP effects	Amino acid substitution	Chemical properties	Interp
4.14731525	G -> G: 0.72863 G -> T: 0.27137	1978	1343	CDS	nonsynonymous	R -> I	basic -> non-polar	
4.14731691	T -> A: 0.29163 T -> T: 0.70837	1812	1276	CDS	nonsynonymous	Y -> N	polar -> polar	
4.14733389	G -> G: 0.5022 G -> T: 0.4978	114	114	CDS	synonymous			
4.14733393	A -> A: 0.50308 A -> C: 0.49692	110	110	CDS	nonsynonymous	Q -> P	polar -> non-polar	
4.14732536	G -> A: 0.50588 G -> G: 0.49412	967	600	CDS	synonymous			
4.14732323	T -> A: 0.13392 T -> T: 0.86608	1180	813	CDS	splice site region, nonsynonymous	N -> K	polar -> basic	
4.14733117	G -> A: 0.01674 G -> G: 0.98326	386	238	CDS	nonsynonymous	V -> I	non-polar -> non-polar	
4.14731132	A -> A: 0.20705 A -> TA: 0.79295	2371	1628	CDS	stop codon, stop codon, frame shift, synonymous			
4.14731188	C -> C: 0.99912 C -> G: 0.00088	2315	1572	CDS	nonsynonymous	N -> K	polar -> basic	
4.14731200	C -> C: 0.99912 C -> T: 0.00088	2303	1560	CDS	synonymous			

If you are interested in further filtering the SNPs in the CDS for, for example, those causing missense mutations, type 'nonsynonymous' into the empty field below the column header 'SNP effects'. Shown are all SNPs which are located in the CDS with an AA substitution effect.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 All [Download](#)

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

SNP_ID (Genomic position)	Base frequencies	Genomic position	MRNA Position	SNP location	SNP effects	Amino acid substitution	Chemical properties	Interpro id
4:14731525	G->G: 0.72863 G->T: 0.27137	1978	1343	CDS	nonsynonymous	R->I	basic -> non-polar	
4:14731691	T->A: 0.29163 T->T: 0.70837	1812	1276	CDS	nonsynonymous	Y->N	polar -> polar	
4:14733593	A->A: 0.50308 A->C: 0.49692	110	110	CDS	nonsynonymous	Q->P	polar -> non-polar	
4:14732323	T->A: 0.13392 T->T: 0.86608	1180	813	CDS	splice site region, nonsynonymous	N->K	polar -> basic	
4:14733117	G->A: 0.01674 G->C: 0.98326	386	238	CDS	nonsynonymous	V->I	non-polar -> non-polar	
4:14731188	G->G: 0.99912 C->G: 0.00088	2315	1572	CDS	nonsynonymous	N->K	polar -> basic	
4:14731216	T->C: 0.00088 T->T: 0.99912	2287	1544	CDS	nonsynonymous	M->T	non-polar -> polar	
4:14731228	G->A: 0.00088 G->G: 0.99912	2275	1532	CDS	nonsynonymous	G->D	non-polar -> acidic	
4:14731234	C->A: 0.00617 C->C: 0.99383	2269	1526	CDS	nonsynonymous	T->N	polar -> polar	
4:14731261	T->C: 0.00088 T->T: 0.99912	2242	1499	CDS	nonsynonymous	I->T	non-polar -> polar	

The first column 'SNP ID' shows all SNPs which are selected with the implemented filter steps by now. The SNP IDs are important to identify accessions which contain SNPs of interest (see chapter 3.1.1.2).

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 All [Download](#)

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

SNP_ID (Genomic position)	Base frequencies	Genomic position	MRNA Position	SNP location	SNP effects	Amino acid substitution	Chemical properties	Interpro id
4:14731525	G->G: 0.72863 G->T: 0.27137	1978	1343	CDS	nonsynonymous	R->I	basic -> non-polar	
4:14731691	T->A: 0.29163 T->T: 0.70837	1812	1276	CDS	nonsynonymous	Y->N	polar -> polar	
4:14733593	A->A: 0.50308 A->C: 0.49692	110	110	CDS	nonsynonymous	Q->P	polar -> non-polar	
4:14732323	T->A: 0.13392 T->T: 0.86608	1180	813	CDS	splice site region, nonsynonymous	N->K	polar -> basic	
4:14733117	G->A: 0.01674 G->C: 0.98326	386	238	CDS	nonsynonymous	V->I	non-polar -> non-polar	
4:14731188	G->G: 0.99912 C->G: 0.00088	2315	1572	CDS	nonsynonymous	N->K	polar -> basic	
4:14731216	T->C: 0.00088 T->T: 0.99912	2287	1544	CDS	nonsynonymous	M->T	non-polar -> polar	
4:14731228	G->A: 0.00088 G->G: 0.99912	2275	1532	CDS	nonsynonymous	G->D	non-polar -> acidic	
4:14731234	C->A: 0.00617 C->C: 0.99383	2269	1526	CDS	nonsynonymous	T->N	polar -> polar	
4:14731261	T->C: 0.00088 T->T: 0.99912	2242	1499	CDS	nonsynonymous	I->T	non-polar -> polar	

Additional interesting information are the nature of AA substitutions and whether the mutated residue may have different physico-chemical properties to specifically focus on such SNPs with a potential consequence on protein structure and/or function.

Results: AT4G30120.1

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

SNP ID (Genomic position)	Base frequencies	Genomic position	MRNA Position	SNP location	SNP effects	Amino acid substitution	Chemical properties	Interpro id
4:14731525	G -> G: 0.72863 G -> T: 0.27137	1978	1343	CDS	nonsynonymous	R -> I	basic -> non-polar	
4:14731691	T -> A: 0.29163 T -> T: 0.70837	1812	1276	CDS	nonsynonymous	Y -> N	polar -> polar	
4:14733390	A -> A: 0.50308 A -> C: 0.49692	110	110	CDS	nonsynonymous	Q -> P	polar -> non-polar	
4:14732323	T -> A: 0.13392 T -> T: 0.86608	1180	813	CDS	splice site region, nonsynonymous	N -> K	polar -> basic	
4:14733117	G -> A: 0.01674 G -> G: 0.98326	386	238	CDS	nonsynonymous	V -> I	non-polar -> non-polar	
4:14731188	C -> C: 0.99912 C -> G: 0.00088	2315	1572	CDS	nonsynonymous	N -> K	polar -> basic	
4:14731216	T -> C: 0.00088 T -> T: 0.99912	2287	1544	CDS	nonsynonymous	M -> T	non-polar -> polar	
4:14731228	G -> A: 0.00088 G -> G: 0.99912	2275	1532	CDS	nonsynonymous	G -> D	non-polar -> acidic	
4:14731234	C -> A: 0.00617 C -> C: 0.99383	2269	1526	CDS	nonsynonymous	T -> N	polar -> polar	
4:14731261	T -> C: 0.00088 T -> T: 0.99912	2242	1499	CDS	nonsynonymous	I -> T	non-polar -> polar	

If you are interested in nonsynonymous SNPs which are located in a protein domain or in combination with already shown information (e.g. physico-chemical property, etc.), scroll to the right to the 'Interpro ID', 'Interpro description', 'Domain ID' and 'Domain description' columns to retrieve this type of information.

Results: AT4G30120.1

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Interpro id	Interpro description	Domain id	Domain description
		PTHHR11939:SF31,PTHHR11939:G3D5A.2.70.150.10	HEAVY METAL CATION TRANSPORT ATPASE (CADMIL)
		PTHHR11939:SF31,PTHHR11939:G3D5A.2.70.150.10	HEAVY METAL CATION TRANSPORT ATPASE (CADMIL)
		PTHHR11939:SF31,PTHHR11939:SSF81660	HEAVY METAL CATION TRANSPORT ATPASE (CADMIL)
		PF00122,1IGR01494,PTHHR11939:SF31,PTHHR11939:SSF81653,SSF81665,G3D5A.2.70.150.10	E1-E2_ATPase_ATPase_P-type: HAD ATPase, P-type, fa
		PTHHR11939:SF31,PTHHR11939:SSF81665,SSF81660	HEAVY METAL CATION TRANSPORT ATPASE (CADMIL)
		PTHHR11939:SF31,PTHHR11939:PS50846,SSF55008	HEAVY METAL CATION TRANSPORT ATPASE (CADMIL)
		PTHHR11939:SF31,PTHHR11939:PS50846,SSF55008	HEAVY METAL CATION TRANSPORT ATPASE (CADMIL)
		PTHHR11939:SF31,PTHHR11939:PS50846,SSF55008	HEAVY METAL CATION TRANSPORT ATPASE (CADMIL)
		PTHHR11939:SF31,PTHHR11939:PS50846,SSF55008	HEAVY METAL CATION TRANSPORT ATPASE (CADMIL)
		PTHHR11939:SF31,PTHHR11939:PS50846,SSF55008	HEAVY METAL CATION TRANSPORT ATPASE (CADMIL)

To download information for the SNPs of interest, generate a table with a format as previously described (see 2.5.1) to identify and download accessions carrying your SNP(s) of interest (see 2.5.2).

Another way to identify SNP IDs for SNPs of interest is to download the complete SNP table (see 2.4.6.6). Open the downloaded file in your data processing software of choice (e.g. Excel, etc.) and filter the columns of interest (SNP location, SNP effects, etc.). The resulting lists can also be used for accession identification (chapter 3.1.1.2).

Here we would like to explain filtering for SNP location and SNP effects. Filter CDS in your calculator program of choice (e.g. Excel) in the column SNP location.

3.1.1.2 Identification of accessions with SNPs of interest

If you have identified a single SNP of interest, you might want to do experiments with accessions carrying different alleles at this locus, or simply clone different alleles from original genetic material. You can click on the ‘SNP ID’ (e.g. 4:14731525) to move directly to the tab ‘SNPs ↔ Accessions’ (see chapter 2.4.3).

Results: AT4G30120.1

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Sequence viewer | **SNPs** | SNPs <-> Accessions | Haplotype <-> Accession table | Protein haplotype <-> Accession table

SNP_ID (Genomic position)	Base frequencies	Genomic position	MRNA Position	SNP location	SNP effects	Amino acid substitution	Chemical properties	Interpro ID
4:14731525	Default: No cutoff	1978	1343	CDS	nonsynonymous	R → I	basic → non-polar	
4:14731691	T → A: 0.29163 T → T: 0.70837	1812	1276	CDS	nonsynonymous	Y → N	polar → polar	
4:14733393	A → A: 0.50308 A → C: 0.49692	110	110	CDS	nonsynonymous	Q → P	polar → non-polar	
4:14732323	T → A: 0.13392 T → T: 0.86608	1180	813	CDS	splice site region, nonsynonymous	N → K	polar → basic	
4:14733117	G → A: 0.01624 G → G: 0.98376	386	238	CDS	nonsynonymous	V → I	non-polar → non-polar	
4:14731188	C → C: 0.99912 C → G: 0.00088	2315	1572	CDS	nonsynonymous	N → K	polar → basic	
4:14731216	T → C: 0.00088 T → T: 0.99912	2287	1544	CDS	nonsynonymous	M → T	non-polar → polar	
4:14731228	G → A: 0.00088 G → G: 0.99912	2275	1532	CDS	nonsynonymous	G → D	non-polar → acidic	
4:14731234	C → A: 0.00517 C → C: 0.99383	2269	1526	CDS	nonsynonymous	T → N	polar → polar	

For the selected SNP of interest all accessions carrying the allele different from the Col-0 reference are now shown in the column ‘Accession_IDs (1001 genomes)’. All shown accessions are different to the reference (Col-0). Now you can search for accessions which are in your lab if you write down the accession ID of interest in the empty field below the column header Accession_IDs (1001 genome).

About

Select transcript and accessions Results for selection Search SNP database

Build & Download GWAS matrices

Results: AT4G30120.1 All Download

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Check selected accessions

Sequence viewer SNPs SNPs <-> Accessions Haplotype <-> Accession table

Protein haplotype <-> Accession table

SNP_ID	Polymorphism	CS numbers and accessions	Accession_IDs (100Tgenomes)
4:14731525			
4:14731525	G -> T	Download CS numbers and accessions	88, 108, 139, 159, 265, 351, 470, 476, 484, 504, 506, 544, 546, 680, 681, 685, 687, 728, 742, 853, 854, 867, 868, 870, 915, 932, 1254, 1257, 1612, 1622, 1651, 1652, 1676, 1684, 1739, 1741, 1756, 1757, 1793, 1797, 1819, 1820, 1834, 1835, 1851, 1852, 1853, 1872, 1942, 1943, 2017, 2031, 2053, 2057, 2081, 2091, 2106, 2108, 2141, 2159, 2166, 2171, 2191, 2212, 2239, 2240, 2276, 2285, 2286, 2370, 2412, 4779, 4826, 4840, 4857, 4884, 4939, 4958, 5023, 5151, 5165, 5210, 5236, 5249, 5253, 5279, 5349, 5353, 5395, 5577, 5644, 5651, 5717, 5718, 5720, 5726, 5741, 5776, 5779, 5798, 5822, 5832, 5836, 6016, 6017, 6024, 6025, 6034, 6038, 6040, 6076, 6087, 6091, 6095, 6096, 6099, 6100, 6104, 6105, 6108, 6113, 6118, 6119, 6126, 6132, 6133, 6134, 6138, 6141, 6145, 6148, 6149, 6150, 6153, 6154, 6163, 6166, 6217, 6284, 6413, 6739, 6740, 6744, 6749, 6750, 6805, 6806, 6814, 6897, 6898, 6907, 6923, 6924, 6926, 6927, 6940, 6943, 6944, 6959, 6966, 6981, 6986, 6989, 6992, 7003, 7008, 7014, 7026, 7028, 7033, 7058, 7061, 7064, 7071, 7094, 7102, 7107, 7109, 7111, 7119, 7143, 7160, 7162, 7164, 7199, 7202, 7208, 7217, 7248, 7250, 7255, 7276, 7298, 7306, 7307, 7314, 7316, 7320, 7332, 7342, 7344, 7358, 7359, 7377, 7378, 7387, 7411, 7415, 7419, 7471, 7475, 7477, 7515, 7521, 7523, 7529, 7530, 7566, 7568, 7757, 7767, 8037, 8057, 8077, 8132, 8171, 8233, 8247, 8312, 8335, 8351, 8366, 8369, 8422, 8464, 8483, 8699, 8723, 9027, 9057, 9298, 9314, 9380, 9395, 9404, 9405, 9407, 9427, 9471, 9532, 9534, 9550, 9559, 9565, 9581, 9584, 9585,

You can download the accession table by clicking on the link in the column 'CS numbers and accessions' (see in chapter 2.4.3).

About

Select transcript and accessions Results for selection Search SNP database

Build & Download GWAS matrices

Results: AT4G30120.1 All Download

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Check selected accessions

Sequence viewer SNPs SNPs <-> Accessions Haplotype <-> Accession table

Protein haplotype <-> Accession table

SNP_ID	Polymorphism	CS numbers and accessions	Accession_IDs (100Tgenomes)
4:14731525			
4:14731525	G -> T	Download CS numbers and accessions	88, 108, 139, 159, 265, 351, 470, 476, 484, 504, 506, 544, 546, 680, 681, 685, 687, 728, 742, 853, 854, 867, 868, 870, 915, 932, 1254, 1257, 1612, 1622, 1651, 1652, 1676, 1684, 1739, 1741, 1756, 1757, 1793, 1797, 1819, 1820, 1834, 1835, 1851, 1852, 1853, 1872, 1942, 1943, 2017, 2031, 2053, 2057, 2081, 2091, 2106, 2108, 2141, 2159, 2166, 2171, 2191, 2212, 2239, 2240, 2276, 2285, 2286, 2370, 2412, 4779, 4826, 4840, 4857, 4884, 4939, 4958, 5023, 5151, 5165, 5210, 5236, 5249, 5253, 5279, 5349, 5353, 5395, 5577, 5644, 5651, 5717, 5718, 5720, 5726, 5741, 5776, 5779, 5798, 5822, 5832, 5836, 6016, 6017, 6024, 6025, 6034, 6038, 6040, 6076, 6087, 6091, 6095, 6096, 6099, 6100, 6104, 6105, 6108, 6113, 6118, 6119, 6126, 6132, 6133, 6134, 6138, 6141, 6145, 6148, 6149, 6150, 6153, 6154, 6163, 6166, 6217, 6284, 6413, 6739, 6740, 6744, 6749, 6750, 6805, 6806, 6814, 6897, 6898, 6907, 6923, 6924, 6926, 6927, 6940, 6943, 6944, 6959, 6966, 6981, 6986, 6989, 6992, 7003, 7008, 7014, 7026, 7028, 7033, 7058, 7061, 7064, 7071, 7094, 7102, 7107, 7109, 7111, 7119, 7143, 7160, 7162, 7164, 7199, 7202, 7208, 7217, 7248, 7250, 7255, 7276, 7298, 7306, 7307, 7314, 7316, 7320, 7332, 7342, 7344, 7358, 7359, 7377, 7378, 7387, 7411, 7415, 7419, 7471, 7475, 7477, 7515, 7521, 7523, 7529, 7530, 7566, 7568, 7757, 7767, 8037, 8057, 8077, 8132, 8171, 8233, 8247, 8312, 8335, 8351, 8366, 8369, 8422, 8464, 8483, 8699, 8723, 9027, 9057, 9298, 9314, 9380, 9395, 9404, 9405, 9407, 9427, 9471, 9532, 9534, 9550, 9559, 9565, 9581, 9584, 9585,

Another way is to move to the 'SNPs ↔ Accessions' tab (see chapter 2.4.3) ...

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 [Download](#)

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

SNP ID (Genomic position)	Base frequencies	Genomic position	MRNA Position	SNP location	SNP effects	Amino acid substitution	Chemical properties	Interpro id
4:14731525	G -> G: 0.72863 G -> T: 0.27137	1978	1343	CDS	nonsynonymous	R -> I	basic -> non-polar	
4:14731691	T -> A: 0.29163 T -> T: 0.70837	1812	1276	CDS	nonsynonymous	Y -> N	polar -> polar	
4:14733390	A -> A: 0.50308 A -> C: 0.49692	110	110	CDS	nonsynonymous	Q -> P	polar -> non-polar	
4:14732323	T -> A: 0.13392 T -> T: 0.86608	1180	813	CDS	splice site region, nonsynonymous	N -> K	polar -> basic	
4:14733117	G -> A: 0.01674 G -> G: 0.98326	386	238	CDS	nonsynonymous	V -> I	non-polar -> non-polar	
4:14731188	C -> C: 0.99912 C -> G: 0.00088	2315	1572	CDS	nonsynonymous	N -> K	polar -> basic	
4:14731216	T -> C: 0.00088 T -> T: 0.99912	2287	1544	CDS	nonsynonymous	M -> T	non-polar -> polar	
4:14731228	G -> A: 0.00088 G -> G: 0.99912	2275	1532	CDS	nonsynonymous	G -> D	non-polar -> acidic	
4:14731234	C -> A: 0.00617 C -> C: 0.99383	2269	1526	CDS	nonsynonymous	T -> N	polar -> polar	
4:14731261	T -> C: 0.00088 T -> T: 0.99912	2242	1499	CDS	nonsynonymous	I -> T	non-polar -> polar	

... and type a single SNP ID of interest (e.g. 4:14731525) into the empty field below the column header 'SNP_ID'.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 [Download](#)

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

SNP_ID	Polymorphism	CS numbers and accessions	Accession_IDs (100Genomes)
4:14731525	G -> T	Download CS numbers and accessions	88, 108, 139, 159, 265, 351, 470, 476, 484, 504, 506, 544, 546, 680, 681, 685, 687, 728, 742, 853, 854, 867, 868, 870, 915, 932, 1254, 1257, 1612, 1622, 1651, 1652, 1676, 1684, 1739, 1741, 1756, 1757, 1793, 1797, 1819, 1820, 1834, 1835, 1851, 1852, 1853, 1872, 1942, 1943, 2017, 2031, 2053, 2057, 2081, 2091, 2106, 2108, 2141, 2159, 2166, 2171, 2191, 2212, 2239, 2240, 2276, 2285, 2286, 2370, 2412, 4779, 4826, 4840, 4857, 4884, 4939, 4958, 5023, 5151, 5165, 5210, 5236, 5249, 5253, 5279, 5349, 5353, 5395, 5577, 5644, 5651, 5717, 5718, 5720, 5726, 5741, 5776, 5779, 5798, 5822, 5832, 5836, 6016, 6017, 6024, 6025, 6034, 6038, 6040, 6076, 6087, 6091, 6095, 6096, 6099, 6100, 6104, 6105, 6108, 6113, 6118, 6119, 6126, 6132, 6133, 6134, 6138, 6141, 6145, 6148, 6149, 6150, 6153, 6154, 6163, 6166, 6217, 6284, 6413, 6739, 6740, 6744, 6749, 6750, 6805, 6806, 6814, 6897, 6898, 6907, 6923, 6924, 6926, 6927, 6940, 6943, 6944, 6959, 6966, 6981, 6986, 6989, 6992, 7003, 7008, 7014, 7026, 7028, 7033, 7058, 7061, 7064, 7071, 7094, 7102, 7107, 7109, 7111, 7119, 7143, 7160, 7162, 7164, 7199, 7202, 7208, 7217, 7248, 7250, 7255, 7276, 7298, 7306, 7307, 7314, 7316, 7320, 7332, 7342, 7344, 7358, 7359, 7377, 7378, 7387, 7411, 7415, 7419, 7471, 7475, 7477, 7515, 7521, 7523, 7529, 7530, 7566, 7568, 7757, 7767, 8037, 8057, 8077, 8132, 8171, 8233, 8247, 8312, 8335, 8351, 8366, 8369, 8422, 8464, 8483, 8699, 8723, 9027, 9057, 9298, 9314, 9380, 9395, 9404, 9405, 9407, 9427, 9471, 9532, 9534, 9550, 9559, 9565, 9581, 9584, 9585,

If you are interested in more than a single SNP, generate an input list with SNP IDs in your data processing software of choice (e.g. Excel) or, alternatively, use a filtered SNP table (see chapter 3.1.1.1). Next, move to the 'Search SNP database' tab followed by 'Get accessions for SNPs' tab.

Results: AT4G30120.1 All Download

ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

Sequence viewer SNPs SNPs <-> Accessions Haplotype <-> Accession table
 Protein haplotype <-> Accession table

SNP_id (Genomic position)	Base frequencies	Genomic position	MRNA Position	SNP location	SNP effects	Amino acid substitution	Chemical properties	Interpro id
4:14731525	G -> G: 0.72863 G -> T: 0.27137	1978	1343	CDS	nonsynonymous	R -> I	basic -> non-polar	
4:14731691	T -> A: 0.29163 T -> T: 0.70837	1812	1276	CDS	nonsynonymous	Y -> N	polar -> polar	
4:14733390	A -> A: 0.50308 A -> C: 0.49692	110	110	CDS	nonsynonymous	Q -> P	polar -> non-polar	
4:14732323	T -> A: 0.13392 T -> T: 0.86608	1180	813	CDS	splice site region, nonsynonymous	N -> K	polar -> basic	
4:14733117	G -> A: 0.01674 G -> G: 0.98326	386	238	CDS	nonsynonymous	V -> I	non-polar -> non-polar	
4:14731188	C -> C: 0.99912 C -> G: 0.00088	2315	1572	CDS	nonsynonymous	N -> K	polar -> basic	
4:14731216	T -> C: 0.00088 T -> T: 0.99912	2287	1544	CDS	nonsynonymous	M -> T	non-polar -> polar	
4:14731228	G -> A: 0.00088 G -> G: 0.99912	2275	1532	CDS	nonsynonymous	G -> D	non-polar -> acidic	
4:14731234	C -> A: 0.00617 C -> C: 0.99383	2269	1526	CDS	nonsynonymous	T -> N	polar -> polar	
4:14731261	T -> C: 0.00088 T -> T: 0.99912	2242	1499	CDS	nonsynonymous	I -> T	non-polar -> polar	

Get SNP information Get accessions for SNPs

Download Accessions

Paste the SNP IDs of interest into the empty field and click 'Download Accessions' to retrieve accession information (see chapter 2.5.2).

3.1.1.3 Identification of accessions with haplotypes of interest

Since SNPs rarely occur in isolation due to phenomena like linkage disequilibrium, you might be interested in identifying haplotypes containing an identical sequence of SNPs for your gene of interest; for example for subsequent cloning projects. We will therefore demonstrate the identification of haplotypes and accessions sharing the same haplotype for a gene of interest.

First, identify SNPs of interest as previously shown (see 3.1.1.2). If you have a list of SNP IDs you can search for a single SNP ID in the tab 'Haplotype ↔ Accession table'. Type your SNP ID of interest (e.g. 4:14731525) into the column 'SNP list' (see chapter 2.4.4).

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 SNPs [Download](#)
 ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3
 Check selected accessions

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Haplotype color	Haplotype_ID	#Accessions	#SNPs	Sequence	SNP list	CS numbers and accessions	Accession_ID
					4:14731925		
	AT4G30120.1.10_dna	1	5	Download Sequence	4:14731925,4:14731691,4:14732176,4:14733044,4:14733376	Download CS numbers and accessions	5832
	AT4G30120.1.11_dna	7	4	Download Sequence	4:14731925,4:14731691,4:14732176,4:14733376	Download CS numbers and accessions	6091,6095,6096
	AT4G30120.1.12_dna	1	3	Download Sequence	4:14731925,4:14731691,4:14733376	Download CS numbers and accessions	6413
	AT4G30120.1.14_dna	1	3	Download Sequence	4:14731925,4:14731691,4:14733044	Download CS numbers and accessions	6105
	AT4G30120.1.16_dna	5	3	Download Sequence	4:14731925,4:14731691,4:14732717	Download CS numbers and accessions	6284,9057,9058
	AT4G30120.1.29_dna	8	3	Download Sequence	4:14731925,4:14731691,4:14732409	Download CS numbers and accessions	915,4840,6540,6541,6542,6543,6544
	AT4G30120.1.35_dna	4	3	Download Sequence	4:14731964,4:14731925,4:14731691	Download CS numbers and accessions	6897,9534,9535,9536

Shown are all haplotypes and the accessions with your SNP of interest.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 SNPs [Download](#)
 ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3
 Check selected accessions

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Haplotype color	Haplotype_ID	#Accessions	#SNPs	Sequence	SNP list	CS numbers and accessions	Accession_ID
					4:14731925		
	AT4G30120.1.10_dna	1	5	Download Sequence	4:14731925,4:14731691,4:14732176,4:14733044,4:14733376	Download CS numbers and accessions	5832
	AT4G30120.1.11_dna	7	4	Download Sequence	4:14731925,4:14731691,4:14732176,4:14733376	Download CS numbers and accessions	6091,6095,6096
	AT4G30120.1.12_dna	1	3	Download Sequence	4:14731925,4:14731691,4:14733376	Download CS numbers and accessions	6413
	AT4G30120.1.14_dna	1	3	Download Sequence	4:14731925,4:14731691,4:14733044	Download CS numbers and accessions	6105
	AT4G30120.1.16_dna	5	3	Download Sequence	4:14731925,4:14731691,4:14732717	Download CS numbers and accessions	6284,9057,9058
	AT4G30120.1.29_dna	8	3	Download Sequence	4:14731925,4:14731691,4:14732409	Download CS numbers and accessions	915,4840,6540,6541,6542,6543,6544
	AT4G30120.1.35_dna	4	3	Download Sequence	4:14731964,4:14731925,4:14731691	Download CS numbers and accessions	6897,9534,9535,9536

If you would like to identify a haplotype which contains only your SNP of interest, write 1 in the search field of column '#SNPs'. Now all haplotypes and accessions which contain only your SNP of interest are shown.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 SNPs [Download](#)
 ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Haplotype color	Haplotype_ID	#Accessions	#SNPs	Sequence	SNP list	CS numbers and accessions	Accession_IDs (1001 genomes)
	AT4G30120.1.42_dna	5	1	<input type="text" value="TT"/>	4:14731525	Download CS numbers and accessions	2171,6744,6966,9941,10022

For our example we identify the haplotype AT4G30120.1.42 with the SNP of interest (4:14731525). Now we can identify accessions which have the haplotype of interest. All accessions are listed in the column 'Accession_IDs'.

[About](#) | [Select transcript and accessions](#) | [Results for selection](#) | [Search SNP database](#)
[Build & Download GWAS matrices](#)

Results: AT4G30120.1 SNPs [Download](#)
 ATHMA3: A. THALIANA HEAVY METAL ATPASE 3; HMA3: Heavy metal atpase 3

[Sequence viewer](#) | [SNPs](#) | [SNPs <-> Accessions](#) | [Haplotype <-> Accession table](#)
[Protein haplotype <-> Accession table](#)

Haplotype color	Haplotype_ID	#Accessions	#SNPs	Sequence	SNP list	CS numbers and accessions	Accession_IDs (1001 genomes)
	AT4G30120.1.42_dna	5	1	<input type="text" value="TT"/>	4:14731525	Download CS numbers and accessions	2171,6744,6966,9941,10022

You can also download a list of the selected accessions.

3.1.1.4 Identification of accessions with protein haplotypes of interest

If you are interested in SNPs which show only AA substitutions then use protein haplotypes. In our example we use the same SNP ID as before (4:14731525) and then move to the tab 'Protein haplotype ↔ Accession table' (see chapter 2.4.5). Filter for SNP ID 4:14731525.

To identify accessions carrying the SNP alleles of interest, follow the same procedure as for haplotype identification described above (see 3.1.1.3).

3.1.2 Identification of known associated SNPs in a gene of interest

In its SNP table SNPstar also lists published associations of SNPs retrieved from the AraGWAS (<https://aragwas.1001genomes.org/#/>) database and, in addition, associations of SNPs with environmental/climatic factors (eGWAS). You can exploit this information to find out whether SNP(s) in your gene of interest are associated with such phenotypes or environmental/climatic data. Select your gene of interest and all accessions (see 2.1).

Move to tab 'SNPs' and check the columns 'AraGWAS top hits', 'GWAS corrected p-values' and 'Number of associated climate factors' (see 2.4.2) to retrieve this information.

3.2 GWAS

3.2.1 GWAS with all transcripts and a subset of accession.

The initial situation here is the availability of phenotypic data for a subset of accessions or the plan to collect phenotypic data for a subset of accessions that are available in your lab. The widest approach to perform a genome wide association study would include all available SNPs. In this case, you can upload a csv file with the subset of accessions, choose "SNP GWAS Matrix" as matrix type and "All SNPs" as SNP subset and compute the GWAS matrix (see chapter 2.6). This will result in a large dataset. For a more reduced dataset, you could choose a subset of SNPs like "coding SNPs" or "nonsynonymous SNPs" which would focus the subsequent analysis on functional SNPs.

As the conventional GWAS exhibits difficulties in detecting rare alleles and is furthermore limited by the mostly biallelic information of the SNPs, the use of "Haplotype GWAS Matrix" or "Proteotype GWAS Matrix" as matrix type is an effective approach to overcome the limitations of biallelic SNPs and increases the allelic resolution of the candidate genomic region (Qian et al. 2017). The proteotype dataset contains fewer variants per transcript compared to the haplotype dataset as only nonsynonymous SNPs that cause an amino acid exchange are considered (see chapter 2.6).

The computed GWAS matrices can be used directly or with some modification as input for different GWAS tools, e.g. TASSEL etc.

3.2.2 GWAS with a subset of transcripts and a subset of accessions

For this use case, the initial situation is again the availability of phenotypic data for a subset of accessions or the plan to collect phenotypic data for a subset of accessions that are available in you lab. Furthermore, you have already focused your interest on a group of transcripts, may it be the large group of transcription factors, auxin signaling genes or a gene family like ARFs (AUXIN RESPONSE FACTORS). After uploading the csv files of your subset of accession and your transcripts of interest, you could compute the "SNP GWAS Matrix" type with "All SNPs" or a "Haplotype GWAS Matrix", if your analysis should contain all SNPs. If you want to focus more on SNPs that cause an amino acid exchange, you can select the "SNP GWAS Matrix" with "nonsynonymous SNPs" or the "Proteotype GWAS Matrix" (see chapter 2.6). The computed GWAS matrices can be used directly or with some modification as input for different GWAS tools, e.g. TASSEL etc.

3.2.3 GWAS with all or a subset of transcripts and all accessions

This would correspond to the rare occasion that phenotypic data is available for all 1135 accessions. If the GWAS matrix should also contain all transcripts you do not need to upload any file, only choose the SNP matrix type and SNP type and compute the GWAS matrix (see also 3.5.1). If you want to focus on a specific subset of transcripts, you upload only the transcript csv file and then compute the GWAS matrix (see also 3.5.2).