

## *ChroMoS* Guide (version 1.2)

### **Background**

Genome-wide association studies (GWAS) reveal increasing number of disease-associated SNPs. Since majority of these SNPs are located in intergenic and intronic regions the assessment of their functionality was hindered by the lack of information about regulatory regions. It requires SNP prioritization for initial analysis to be followed by more focused functional analysis.

ChroMoS (Chromatin Modified SNPs) combines genetic and epigenetic data with the goal to facilitate SNP classification and prioritization. To this end the user can provide SNP data in VCF format, dbSNPs or select GWAS SNPs from the local database. The user provides annotations for chromatin state regions obtained from pre-calculated segmentation of epigenomic data for ENCODE 9 cell types. The genome segmentation based on chromatin marks allows predictions of functional elements, such as enhancers and promoters. In fact, six major categories of chromatin states were distinguished: enhancer, insulator, transcribed, repressed and inactive states. Promoter category was further partitioned into 3 states: active, weak and poised based on the expression level of adjacent genes; enhancer class was segregated into strong and weak states. Transcribed regions were separated into strongly and weakly transcribed regions. Also, heterochromatic and repetitive states were isolated based on their H3K9me3 enrichment. Polycomb-repressed regions were defined as well. In total, 15 states were distinguished and this data has been used in ChroMoS. It was shown that disease-associated SNPs were more likely to be situated within strong enhancer regions than neutral dbSNPs. Particularly, it was evident for cell types related to a disease, e.g. lymphoblastoid cell (GM12878) enhancers contained SNPs associated with systemic lupus erythematosus [Ernst et al. (2011), *Nature*].

Based on this data ChroMoS suggests the functional impact of a SNP. In the process, SNPs are assigned to the various chromatin states. The chromatin states were computed applying multivariate hidden Markov model [Ernst et al. (2011), *Nature*]. It uses patterns of chromatin marks to reduce large combinatorial space to an interpretable set of chromatin states. SNPs positioned in enhancer or transcription states can be subjected to differential analysis of transcription factor binding with sTRAP, and SNPs with potential impact on post-transcriptional

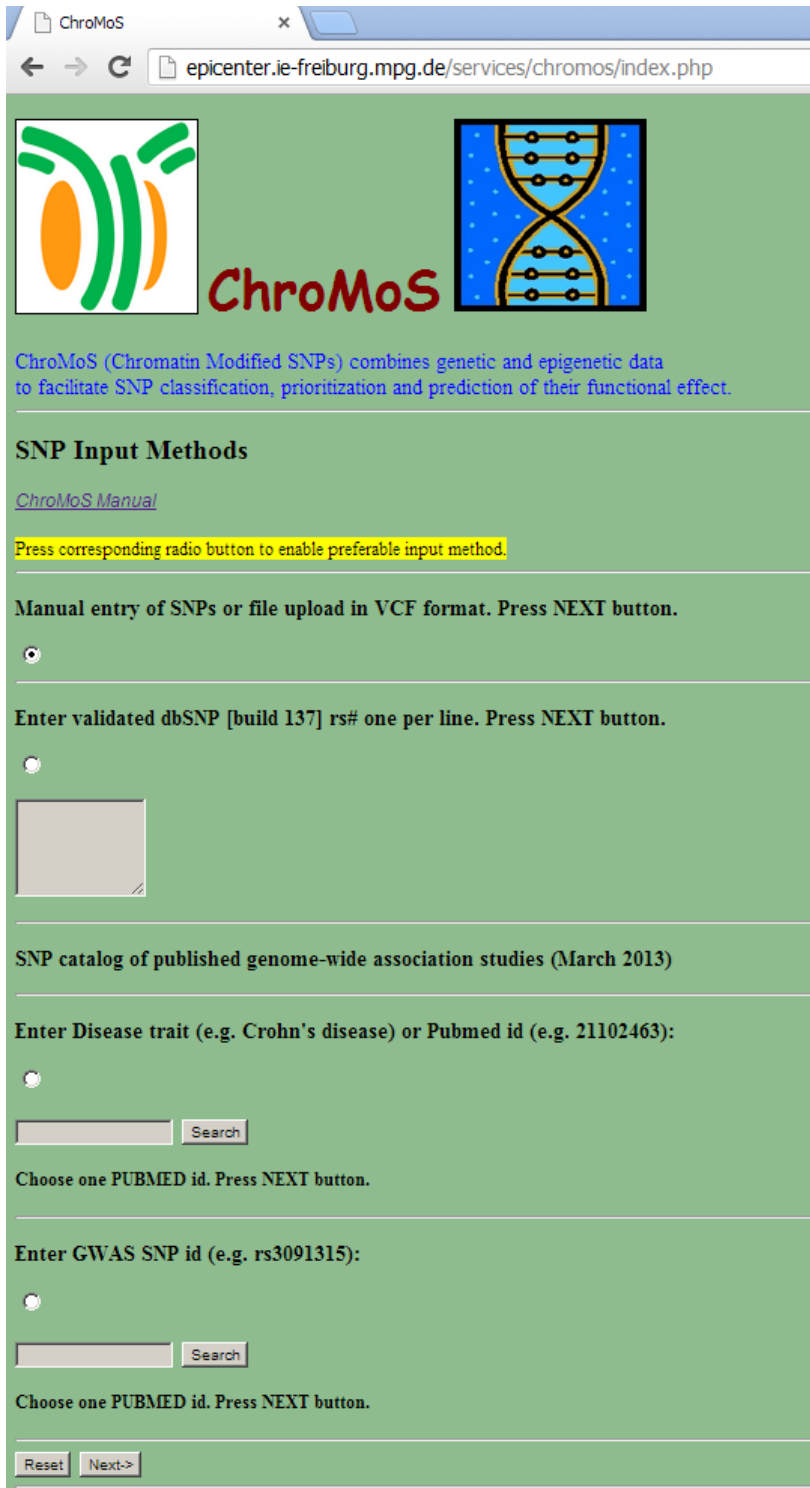
mechanisms are evaluated by MicroSNiPer for a differential binding capacity of annotated miRNA.

sTRAP, analyzes variations in the DNA sequence and predicts quantitative changes to the binding strength of any transcription factor for which there is a binding model. It suggests possible consequences of sequence variations on regulatory networks. The method was tested against a set of known associations between SNPs and their regulatory effects. Its predictions are robust with respect to different parameters and model assumptions. This tool can serve as important point for routine analysis of disease-associated sequence regions [Manke et al. (2010) Hum Mutat].

MicroSNiPer predicts the impact of a SNP on putative microRNA targets. This application interrogates the 3'-untranslated region and predicts if a SNP within the target site will disrupt/eliminate or enhance/create a microRNA binding site. MicroSNiPer computes these sites and examines the effects of SNPs in real time. It has straightforward graphical representation of the results [Barenboim et al. (2010) Hum Mutat].

# ChroMoS Manual

*Warning: Firefox web-browser might not display properly a color map of more than 1000 SNPs.  
Download the map through the web-link.*



The screenshot shows the ChroMoS web interface in a browser window. The address bar shows the URL: `epicenter.ie-freiburg.mpg.de/services/chromos/index.php`. The page has a green background and features the ChroMoS logo (a stylized orange and green 'C' shape) and a DNA double helix icon. Below the logo, the text reads: "ChroMoS (Chromatin Modified SNPs) combines genetic and epigenetic data to facilitate SNP classification, prioritization and prediction of their functional effect." The main section is titled "SNP Input Methods" and includes a link to the "ChroMoS Manual". A yellow highlight indicates the instruction: "Press corresponding radio button to enable preferable input method." There are four input methods, each with a radio button and a description: 1. "Manual entry of SNPs or file upload in VCF format. Press NEXT button." 2. "Enter validated dbSNP [build 137] rs# one per line. Press NEXT button." 3. "SNP catalog of published genome-wide association studies (March 2013)" with a sub-instruction "Enter Disease trait (e.g. Crohn's disease) or Pubmed id (e.g. 21102463):" and a search box. 4. "Enter GWAS SNP id (e.g. rs3091315):" with a search box. At the bottom, there are "Reset" and "Next->" buttons.

ChroMoS

ChroMoS (Chromatin Modified SNPs) combines genetic and epigenetic data to facilitate SNP classification, prioritization and prediction of their functional effect.

### SNP Input Methods

[ChroMoS Manual](#)

Press corresponding radio button to enable preferable input method.

☐ Manual entry of SNPs or file upload in VCF format. Press NEXT button.

☐ Enter validated dbSNP [build 137] rs# one per line. Press NEXT button.

☐ SNP catalog of published genome-wide association studies (March 2013)

Enter Disease trait (e.g. Crohn's disease) or Pubmed id (e.g. 21102463):

Choose one PUBMED id. Press NEXT button.

☐ Enter GWAS SNP id (e.g. rs3091315):

Choose one PUBMED id. Press NEXT button.

*The first page* of Chromos allows **four** input methods. To be able to activate each method a user has to **press corresponding radio button first**. Manual entry is default.

(1) Manual entry of SNPs on the following page. A user simply presses **Next** button on the bottom of the page.

A user is simply directed to the following page where she can upload SNP file in VCF or paste data in VCF into the text field.

Enter validated dbSNP rs# (one per line). Press NEXT button.

☐

rs11134178	▲
rs2157697	
rs6501530	
rs12301774	▼
rs2594278	↕

---

Enter Disease trait (e.g. Crohn's disease) or Pubmed id (e.g. 21102463):

☐

Choose one PUBMED id. Press NEXT button.

---

Enter GWAS SNP id (e.g. rs3091315):

☐

Choose one PUBMED id. Press NEXT button.

(2) Entry of validated dbSNP rs# (~45 mln dbSNPs). One rs# per line. It can be any dbSNP not necessarily from GWAS catalog. After pressing **Next** button these SNPs appear in the second page SNP area in VCF.

Enter Disease trait (e.g. Crohn's disease) or Pubmed id (e.g. 21102463):

☐

Choose one PUBMED id. Press NEXT button.

---

Enter GWAS SNP id (e.g. rs3091315):

☐

Choose one PUBMED id. Press NEXT button.

[Bioinformatics and Next Generation Sequencing Group](#); Max Planck Institute of Immunobiology and Epigenetics

(3) Entry by disease trait (e.g. Crohn's disease) or Pubmed id (e.g. 21102463). Click second from the top radio button. Enter **Crohn's disease**. Click **Search** button. It retrieves a list of all currently published Crohn's disease GWAS studies including unique Pubmed IDs.

Choose a certain Pubmed id e.g. **21102463**. Press **Next** button. ChroMoS retrieves all 71 SNPs belonging to GWA study with PMID 21102463 and displays them on the next page.

---

Enter Disease trait (e.g. Crohn's disease) or Pubmed id (e.g. 21102463):

☐

Choose one PUBMED id. Press NEXT button.

---

Enter GWAS SNP id (e.g. rs3091315):

☐

Choose one PUBMED id. Press NEXT button.

---

---

[Bioinformatics and Next Generation Sequencing Group; Max Planck Institute of Immunobiology and Epigenetics](#)

(4) A user can also retrieve PMID by entering SNP id (e.g. rs3091315) and after pressing **Search** button choose proper PMID. Press **Next** button. Chromos retrieves all 71 SNPs belonging to GWA study with PMID: 21102463 and displays them on the next page.

Warning: Firefox web-browser might not display properly a color map of more than 1000 SNPs. Download the

Enter SNPs in VCF (Variant Call Format):

☒ VCF

```
# PubmedID=21102463 Crohn's disease
chr1 7879063 rs2797685 C T . . .
chr1 67705958 rs11209026 G A . . .
chr1 114377568 rs2476601 A G . . .
chr1 155230131 rs1142287 C T . . .
chr1 160830268 rs4656940 A G . . .
chr1 172853460 rs7517810 C T . . .
chr1 197727642 rs1998598 A G . . .
chr1 200877562 rs7554511 C A . . .
chr1 206939904 rs3024505 G A . . .
```

Upload file with SNPs in VCF (file limit 1000 Kb):

No file chosen

[1000 SNP VCF test file](#)

Choose cell types with pre-computed chromatin states [Ernst et al. (2011) Nature]:

Gm12878H1  
H1hescHMM  
Hepg2HMM  
HmecHMM  
HsmmHMM  
HuvecHMM  
K562HMM  
NhokHMM  
NhiHMM

[Bioinformatics and Next Generation Sequencing Group, Max Planck Institute of Immunobiology and Epigenetics](#)

**On the second page** 71 SNPs from GWA study with PMID 21102463 are displayed in VCF. On this stage user can add her own data by entering her data in the same format. If a user wants to upload only her own SNP file in VCF she can use **Choose File button**. In this case all data in VCF text area are erased. Pressing **Reset button** will recover original data. We provide a test file of 1,000 SNPs in VCF. It can be pasted to VCF area or uploaded as a VCF file directly from the local computer.

**Important:** one SNP record has to be in one continuous line. If this is not a case, text field should be stretched by grabbing lower right corner of the VCF text area.

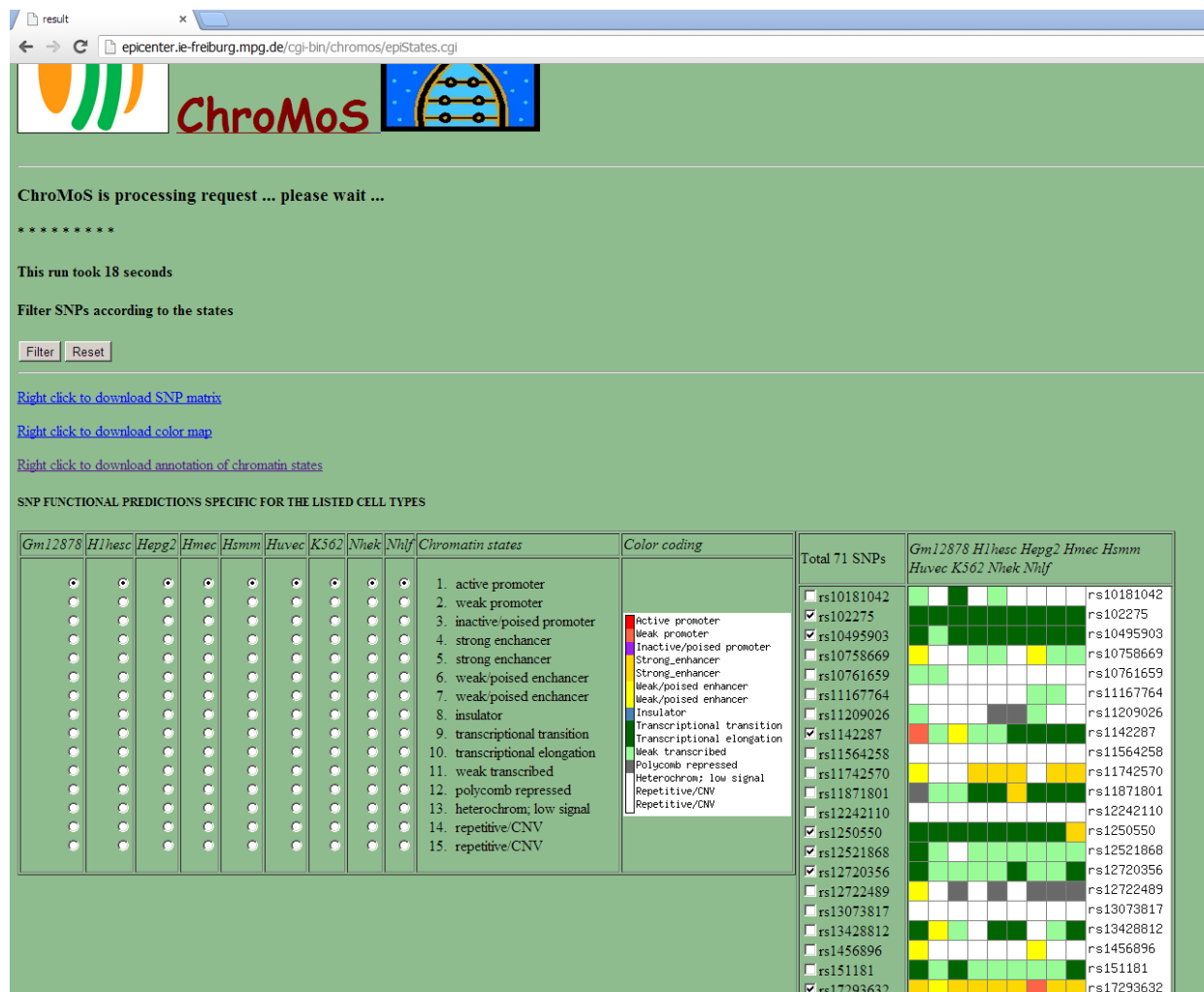
A user can select one or more available cell types with pre-computed chromatin states

in bed-format [Ernst et al. (2011), *Nature*] by Ctrl-Click and press **Run Chromos button**. This invokes Perl CGI script which utilizes bedtools [Quinlan and Hall (2010), *Bioinformatics*] intersecting SNP coordinates with coordinates of chromatin states and, subsequently, matrix2png [Pavlidis and Noble (2003), *Bioinformatics*], which provides color map of 15 states for each cell type.

On *ChroMoS* result page a user can also download digital matrix based on which color map is created and use in other tools. Table includes color map with SNP id aligned to color code of chromatin states. Column names display a number of SNPs and chosen cell types. **Warning:** **Firefox web-browser** has some limitation on displaying large PNG files (above ~ 1,000 SNPs) and alignment for large files is not exact, too. **Opera web-browser** has also graphical limitations.

Next, a user should decide which way she prefers to filter results. One option is to use radio buttons in order to create certain pattern of states, e.g. “active promoter” in all 9 cell types. It is helpful for large SNP sets with only several cell types, or else, this type of selection likely produces empty set. Currently, the limitation for upload is 10,000 SNPs. If the SNP set consists of only several hundred SNPs, we suggest visually examining color map and manually checking out SNPs of interests (e.g. SNPs in the enhancer state in all 9 cell types).

If user starts manually checking out SNPs, pattern filtering is disabled. In order to return to pattern filtering and clear checkboxes user has to press **Reset button**. In this example 11 SNPs were checked out, and then **Filter button** was pressed.

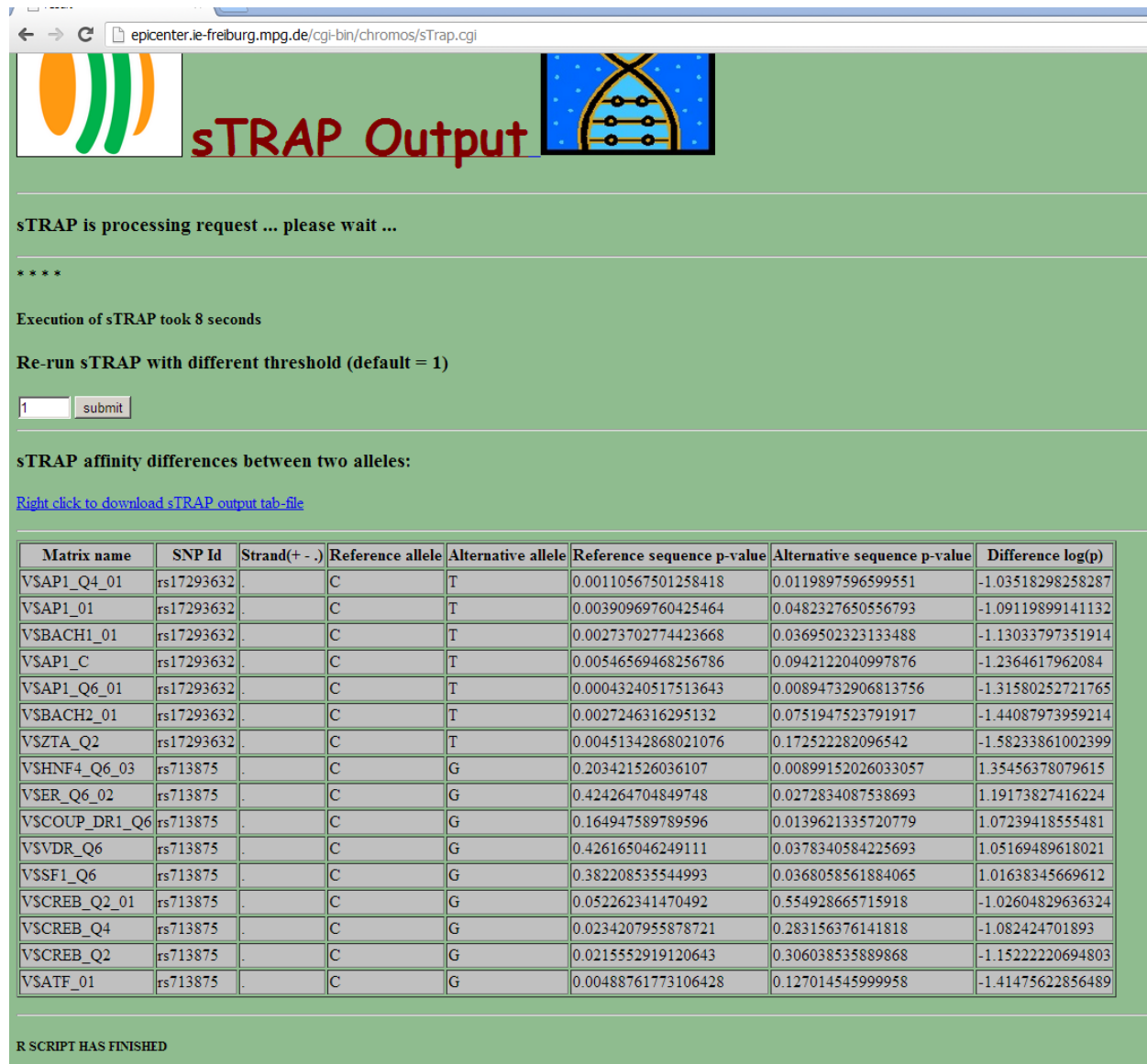


The screenshot shows the ChroMoS web interface. At the top, there's a header with the ChroMoS logo and a navigation bar. Below the header, a message states "ChroMoS is processing request ... please wait ...". A status bar indicates "This run took 18 seconds". A section titled "Filter SNPs according to the states" contains a "Filter" button and a "Reset" button. Below this, there are links to download the SNP matrix, color map, and chromatin states. A section titled "SNP FUNCTIONAL PREDICTIONS SPECIFIC FOR THE LISTED CELL TYPES" displays a table with columns for cell types (Gm12878, H1hesc, Hepg2, Hmec, Hsmm, Huvec, K562, Nhek, Nhlf) and a list of 15 chromatin states. A color coding legend is provided, mapping colors to specific chromatin states. To the right, a table shows the results for 71 SNPs, with a color-coded matrix indicating the state of each SNP across the cell types. The matrix is color-coded according to the legend, with green representing active promoters, yellow representing weak promoters, and red representing inactive/poised promoters.

Cell Type	Gm12878	H1hesc	Hepg2	Hmec	Hsmm	Huvec	K562	Nhek	Nhlf	Chromatin states	Color coding	Total 71 SNPs	Cell Type	Gm12878	H1hesc	Hepg2	Hmec	Hsmm	Huvec	K562	Nhek	Nhlf		
1.	active promoter												rs10181042											rs10181042
2.	weak promoter												rs102275											rs102275
3.	inactive/poised promoter												rs10495903											rs10495903
4.	strong enhancer												rs10758669											rs10758669
5.	strong enhancer												rs10761659											rs10761659
6.	weak/poised enhancer												rs11167764											rs11167764
7.	weak/poised enhancer												rs11209026											rs11209026
8.	insulator												rs1142287											rs1142287
9.	transcriptional transition												rs11564258											rs11564258
10.	transcriptional elongation												rs11742570											rs11742570
11.	weak transcribed												rs11871801											rs11871801
12.	polycomb repressed												rs12242110											rs12242110
13.	heterochrom; low signal												rs1250550											rs1250550
14.	repetitive/CNV												rs12521868											rs12521868
15.	repetitive/CNV												rs12720356											rs12720356
													rs12722489											rs12722489
													rs13073817											rs13073817
													rs13428812											rs13428812
													rs1456896											rs1456896
													rs151181											rs151181
													rs17293632											rs17293632

On the next page filtered SNPs with color code are displayed. Then, in order to test if SNPs affect transcription factor binding a user can send SNPs to sTRAP [Manke et al., (2010) Hum Mutat.] selecting SNPs and pressing **Submit button**. Since sTRAP is computationally intensive, there is a limit of 60 SNPs to submit to sTRAP. Initial threshold is equal to one which displays

only significant candidate SNPs for impact on transcription factor binding sites. However, if there is **an empty result table** a user can decrease threshold (e.g. 0.6) and re-run sTRAP.



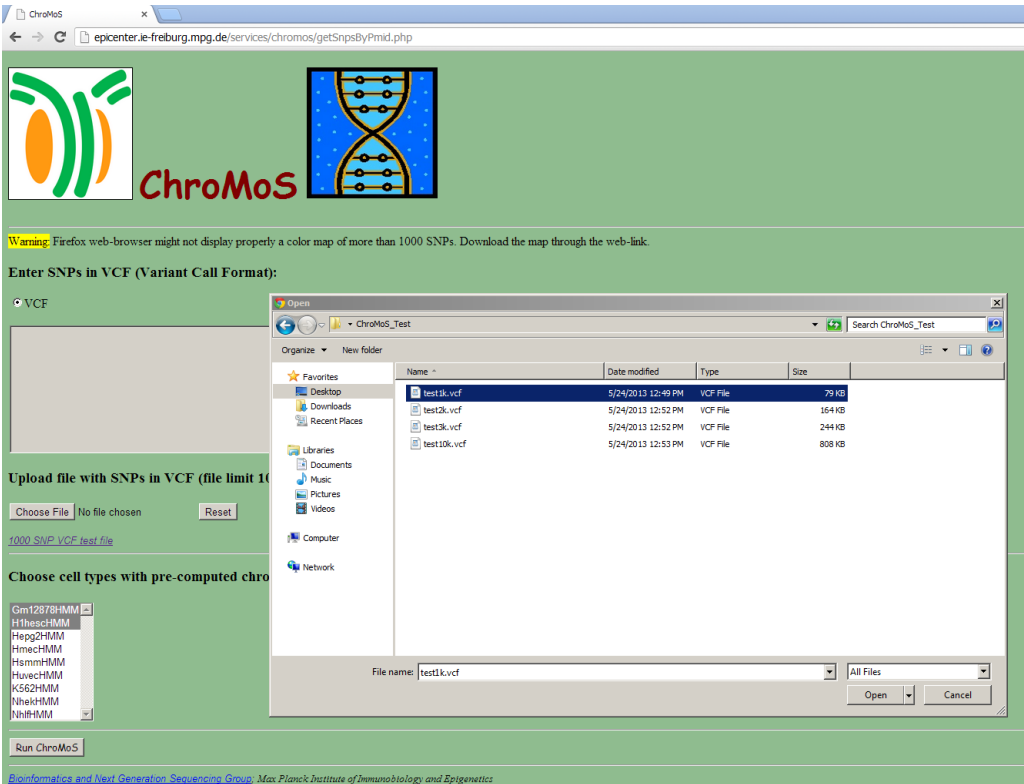
The image shows a screenshot of the sTRAP web interface. At the top, there is a header with the text "sTRAP Output" and a logo. Below the header, a message states "sTRAP is processing request ... please wait ...". A progress bar shows the execution of sTRAP took 8 seconds. A section titled "Re-run sTRAP with different threshold (default = 1)" includes a text input field with the value "1" and a "submit" button. Below this, a message says "sTRAP affinity differences between two alleles:" followed by a link "Right click to download sTRAP output tab-file". The main part of the interface is a table with 8 columns: Matrix name, SNP Id, Strand(+ -), Reference allele, Alternative allele, Reference sequence p-value, Alternative sequence p-value, and Difference log(p). The table contains 20 rows of data. At the bottom, a message states "R SCRIPT HAS FINISHED".

Matrix name	SNP Id	Strand(+ -)	Reference allele	Alternative allele	Reference sequence p-value	Alternative sequence p-value	Difference log(p)
V\$AP1_Q4_01	rs17293632	.	C	T	0.00110567501258418	0.0119897596599551	-1.03518298258287
V\$AP1_01	rs17293632	.	C	T	0.00390969760425464	0.0482327650556793	-1.09119899141132
V\$BACH1_01	rs17293632	.	C	T	0.00273702774423668	0.0369502323133488	-1.13033797351914
V\$AP1_C	rs17293632	.	C	T	0.00546569468256786	0.0942122040997876	-1.2364617962084
V\$AP1_Q6_01	rs17293632	.	C	T	0.00043240517513643	0.00894732906813756	-1.31580252721765
V\$BACH2_01	rs17293632	.	C	T	0.0027246316295132	0.0751947523791917	-1.44087973959214
V\$ZTA_Q2	rs17293632	.	C	T	0.00451342868021076	0.172522282096542	-1.58233861002399
V\$HNF4_Q6_03	rs713875	.	C	G	0.203421526036107	0.00899152026033057	1.35456378079615
V\$SER_Q6_02	rs713875	.	C	G	0.424264704849748	0.0272834087538693	1.19173827416224
V\$COUP_DR1_Q6	rs713875	.	C	G	0.164947589789596	0.0139621335720779	1.07239418555481
V\$VDR_Q6	rs713875	.	C	G	0.426165046249111	0.0378340584225693	1.05169489618021
V\$SF1_Q6	rs713875	.	C	G	0.382208535544993	0.0368058561884065	1.01638345669612
V\$CREB_Q2_01	rs713875	.	C	G	0.052262341470492	0.554928665715918	-1.02604829636324
V\$CREB_Q4	rs713875	.	C	G	0.0234207955878721	0.283156376141818	-1.082424701893
V\$CREB_Q2	rs713875	.	C	G	0.0215552919120643	0.306038535889868	-1.15222220694803
V\$ATF_01	rs713875	.	C	G	0.00488761773106428	0.127014545999958	-1.41475622856489

The sTRAP result page will display transfac matrix names grouped by SNPs. The transcription factors with reduced affinity receive a negative ratio of p-values and those with increased binding get a positive ratio. On the sTRAP result page user can re-run sTRAP with a different threshold. On each step a user can download data in tab-format.



To demonstrate integration with MicroSNiPer [Barenboim *et al.* (2010) *Hum Mutat*], we download 1,000 SNPs sample file with **Choose File button**. We select two cell types GM12878 and H1hesc and press **Run ChroMoS button**.



ChroMoS is processing request ... please wait ...

\*\*

This run took 4 seconds

Filter SNPs according to the states

Filter Reset

[Right click to download SNP matrix](#)


[Right click to download color map](#)

[Right click to download annotation of chromatin states](#)

SNP FUNCTIONAL PREDICTIONS SPECIFIC FOR THE LISTED CELL TYPES

Gm12878	H1hesc	Chromatin states	Color coding	Total 1000 SNPs	Gm12878 H1hesc
<input type="radio"/>	<input type="radio"/>	1. active promoter		<input type="checkbox"/> rs10159187	<input type="checkbox"/> rs10159187
<input type="radio"/>	<input type="radio"/>	2. weak promoter		<input type="checkbox"/> rs10449222	<input type="checkbox"/> rs10449222
<input type="radio"/>	<input type="radio"/>	3. inactive/poised promoter		<input type="checkbox"/> rs10449893	<input type="checkbox"/> rs10449893
<input type="radio"/>	<input type="radio"/>	4. strong enhancer		<input type="checkbox"/> rs1046878	<input type="checkbox"/> rs1046878
<input type="radio"/>	<input type="radio"/>	5. strong enhancer		<input type="checkbox"/> rs1048488	<input type="checkbox"/> rs1048488
<input type="radio"/>	<input type="radio"/>	6. weak/poised enhancer		<input type="checkbox"/> rs10489588	<input type="checkbox"/> rs10489588
<input type="radio"/>	<input type="radio"/>	7. weak/poised enhancer		<input type="checkbox"/> rs10489589	<input type="checkbox"/> rs10489589
<input type="radio"/>	<input type="radio"/>	8. insulator		<input type="checkbox"/> rs10492936	<input type="checkbox"/> rs10492936
<input type="radio"/>	<input type="radio"/>	9. transcriptional transition		<input type="checkbox"/> rs10492940	<input type="checkbox"/> rs10492940
<input type="radio"/>	<input type="radio"/>	10. transcriptional elongation		<input type="checkbox"/> rs10492941	<input type="checkbox"/> rs10492941
<input type="radio"/>	<input type="radio"/>	11. weak transcribed		<input type="checkbox"/> rs10737190	<input type="checkbox"/> rs10737190
<input type="radio"/>	<input type="radio"/>	12. polycomb repressed		<input type="checkbox"/> rs10752733	<input type="checkbox"/> rs10752733
<input type="radio"/>	<input type="radio"/>	13. heterochrom; low signal		<input type="checkbox"/> rs10752737	<input type="checkbox"/> rs10752737
<input type="radio"/>	<input type="radio"/>	14. repetitive/CNV		<input type="checkbox"/> rs10752741	<input type="checkbox"/> rs10752741
<input type="radio"/>	<input type="radio"/>	15. repetitive/CNV		<input type="checkbox"/> rs10752748	<input type="checkbox"/> rs10752748
				<input type="checkbox"/> rs10797342	<input type="checkbox"/> rs10797342
				<input type="checkbox"/> rs10797368	<input type="checkbox"/> rs10797368
				<input type="checkbox"/> rs10797380	<input type="checkbox"/> rs10797380
				<input type="checkbox"/> rs10797384	<input type="checkbox"/> rs10797384

On ChroMoS result page we choose out of 1,000 SNPs all SNPs which are in *transcriptional elongation* state by pressing **radio button** **pattern filtering**. Pressing **Filter button** will bring another page.



Find SNPs affecting Transcription Factor binding with [sTRAP](#) (select no more than 55)

or predict SNP impact on microRNA target sites with [MicroSNiPer](#)

MicroSNiPer

[Right click to download SNP matrix](#)

[Right click to download color map](#)

[Right click to download annotation of chromatin states](#)

SNP FUNCTIONAL PREDICTIONS SPECIFIC FOR THE LISTED CELL TYPES

Color coding	Total 54 SNPs	Gm12878 H1hesc
Active promoter	<input checked="" type="checkbox"/> rs1107910	rs1107910
Weak promoter	<input checked="" type="checkbox"/> rs11260570	rs11260570
Inactive/poised promoter	<input checked="" type="checkbox"/> rs11260611	rs11260611
Strong enhancer	<input checked="" type="checkbox"/> rs1129332	rs1129332
Weak/poised enhancer	<input checked="" type="checkbox"/> rs1129333	rs1129333
Insulator	<input checked="" type="checkbox"/> rs1129333	rs1129333
Transcriptional transition	<input checked="" type="checkbox"/> rs1153105	rs1153105
Transcriptional elongation	<input checked="" type="checkbox"/> rs12032637	rs12032637
Weak transcribed	<input checked="" type="checkbox"/> rs12069909	rs12069909
Polycomb repressed	<input checked="" type="checkbox"/> rs12103	rs12103
Heterochrom: low signal	<input checked="" type="checkbox"/> rs12142199	rs12142199
Repetitive/OW	<input checked="" type="checkbox"/> rs12402622	rs12402622
Repetitive/OW	<input checked="" type="checkbox"/> rs12563338	rs12563338
Repetitive/OW	<input checked="" type="checkbox"/> rs12742323	rs12742323
Repetitive/OW	<input checked="" type="checkbox"/> rs13303287	rs13303287
Repetitive/OW	<input checked="" type="checkbox"/> rs16825336	rs16825336
Repetitive/OW	<input checked="" type="checkbox"/> rs2180311	rs2180311
Repetitive/OW	<input checked="" type="checkbox"/> rs2294489	rs2294489
Repetitive/OW	<input checked="" type="checkbox"/> rs2296715	rs2296715
Repetitive/OW	<input checked="" type="checkbox"/> rs2296716	rs2296716
Repetitive/OW	<input checked="" type="checkbox"/> rs2340582	rs2340582
Repetitive/OW	<input checked="" type="checkbox"/> rs2645081	rs2645081
Repetitive/OW	<input checked="" type="checkbox"/> rs2840532	rs2840532
Repetitive/OW	<input checked="" type="checkbox"/> rs2862157	rs2862157
Repetitive/OW	<input checked="" type="checkbox"/> rs3001336	rs3001336
Repetitive/OW	<input checked="" type="checkbox"/> rs3001344	rs3001344
Repetitive/OW	<input checked="" type="checkbox"/> rs3122920	rs3122920




## ChroMoS to MicroSNiPer

Choose Set of SNPs (dbSNP build 137):

validated dbSNPs

Tab-delimited ChroMoS\_SNPs: chr position SNP\_id ref\_allele alt\_allele (e.g. chr1 9832359 rs7415181 G A)

chr1	1692321	rs1107910	T	C
chr1	1201640	rs11260570	T	C
chr1	1498377	rs11260611	C	T
chr1	2336210	rs1129332	C	T
chr1	2335676	rs1129333	A	G
chr1	1415099	rs1153105	T	C
chr1	1465382	rs12032637	A	G
chr1	1727391	rs12069909	G	A
chr1	1247494	rs12103	A,G	
chr1	1249187	rs12142199	G	A

SNPs in 3'UTRs (hg19 assembly) (select only one SNP)

<input type="radio"/> rs1129332	chr1:2336210	C/T
<input checked="" type="radio"/> rs1129333	chr1:2335676	A/G
<input type="radio"/> rs3818448	chr1:1246972	C/T
<input type="radio"/> rs6659884	chr1:1717848	G/T

Choose one RefSeq id. Press NEXT button.

RER1:NM\_007033

Bioinformatics and Next Generation Sequencing Group, Max Planck Institute of Immunobiology and Epigenetics

On the filter result page there are 54 SNPs which are in *transcriptional elongation* state in both cell types. There is a possibility that some of them are in 3'UTR and can have an impact on microRNA target sites. In order to send these SNPs to integrated tool Microsniper a user has to choose MicroSNiPer from a menu on the top of the page. All SNPs will be automatically checked out. By pressing **Submit button** user send them to *ChroMoS to MicroSNiPer* page. On this stage a user can also add her SNPs in suggested format. Then, user tests if some of these SNPs are in 3'UTRs of RefSeq genes by pressing **Find SNPs in 3'UTRs button**.

MicroSNiPer

epicenter.ie-freiburg.mpg.de/services/microsniper/getSeqByNM.php

**Pick MicroRNA set (miRBase Release 19: Aug 2012)**

hsa microRNAs

**Enter 3'UTR nucleotide sequence in FASTA format (MAX 10 kb):**

```
>NM_007033 range=chr1:2334564-2336885 strand=+
AAGCGGGAGCTGAGGCTGCTCAGCTGTTGCAAGAACAGTTTGTAGCCATTGTTAACAATGCTTTTCTTCAC
ATAAAGTAGTTGATTACGAGGAGTCAATTTTCTTTTAAAAAGAGCTTCAATGATTGTAACTGAAATATC
AGGTTCTAGAGAACTGAGGCTTAAACAAATCCAGTGAATTTCTTTTCACTGACGTTCAAGTGTCTTCAC
GATGGAATTTAGTCACTCAGGCGGGAAGCCAGGCGGTTGAGCCCATGGAGCAAGGCGAGTGGCCGCT
CCCCGCTGTGCCAGTGGGACGAGGAGCAAGGCTGCGAGGAGGAGAACGGCCGCTCCCCGCCAGCCGCTT
CCCCAGCAGCGCAGTGTGTGCCAGCACTCCACAGAGCCGAGGAGTGAATCTAGCTGATTCTGCTGCTGCTG
```

Choose File No file chosen Reset

**Enter SNP id, position relative to the start of the above sequence, allele1/allele2 (e**

rs144802592, 108, C/G  
rs76184090, 160, A/G  
rs188820327, 164, C/T  
rs148530549, 244, C/T

Update SNP List Reset

press button Update List before proceeding further

**SNP List (select at least one SNP)**

<input checked="" type="checkbox"/>	rs144802592	108	C/G
<input type="checkbox"/>	rs76184090	160	A/G
<input checked="" type="checkbox"/>	rs188820327	164	C/T
<input type="checkbox"/>	rs148530549	244	C/T
<input type="checkbox"/>	rs41315648	293	C/T
<input type="checkbox"/>	rs115090288	300	C/T
<input type="checkbox"/>	rs77592022	512	G/T
<input type="checkbox"/>	rs181572880	817	C/T
<input type="checkbox"/>	rs116414456	852	A/G
<input type="checkbox"/>	rs1129333	1113	G/A
<input type="checkbox"/>	rs185473832	1160	A/C
<input type="checkbox"/>	rs112157120	1165	C/T

MicroSNiPer

**Specify minimum seed length:**

6-mer submit

**OUTPUT SPECIFIC TO**

**POLYMORPHIC VARIANT: T[164]-C[108]**

Gene: NM_007033 range=chr1:2334564-2336885 strand=+	SNP: rs188820327 164[C/T]	MicroRNA set: hsa microRNAs Seed length: 8 bp
<pre>190 180 170 160 150 140 AAGAAATCCATGCGATTGTTTAAAGCGCCAAATTCTTCTAGAACCTGATATTTCAGTTA : : : : : AAAAUUUUUUUACUACUACUAG 10 20</pre>		
<a href="#">hsa-miR-3606-3p MIMAT0022965 Homo sapiens miR-3606-3p (21 nt)</a>		

**POLYMORPHIC VARIANT: C[164]-G[108]**

Gene: NM_007033 range=chr1:2334564-2336885 strand=+	SNP: rs144802592 108[C/G]	MicroRNA set: hsa microRNAs Seed length: 7 bp
<pre>140 130 120 110 100 90 CAGTTACAAATCATTGAAGCTCCTTTTAAAAAATAAATTTGACTCCCTCGTAATCAACT : : : : : AAACAAACAUGGUGACUUCUU 10 20</pre>		
<a href="#">hsa-miR-495-3p MIMAT0002817 Homo sapiens miR-495-3p (22 nt)</a>		

Program filters SNPs for presenting in 3'UTRs and creates a table with radio buttons. User has to choose a single SNP from the table, and subsequently a transcript NM\_id from the dropdown list. User also can choose validated dbSNPs (default) or a set of HapMap SNPs on the top of the page. Pressing **Next button** inputs this data to a routine MicroSNiPer workflow. A SNP selected with radio button is added to the list of validated dbSNPs (or HapMap SNPs) positioned within chosen 3'UTR. On MicroSNiPer page a user can also add her own SNPs. Then, user presses **Update SNP List button**, check out SNPs of interest (limit 6 SNPs) and presses **Run Microsniper button**. User can also go directly to MicroSNiPer main page on the <http://epicenter.ie-freiburg.mpg.de/services/microsniper/>.

## REFERENCES

- Barenboim, M. et al. (2010) MicroSNiPer: a web tool for prediction of SNP effects on putative microRNA targets, *Hum Mutat*, 31, 1223-1232.
- Ernst, J. et al. (2011) Mapping and analysis of chromatin state dynamics in nine human cell types, *Nature*, 473, 43-49.
- Franke, A. et al. (2010) Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci, *Nat Genet*, 42, 1118-1125.
- Manke, T. et al. (2010) Quantifying the effect of sequence variation on regulatory interactions, *Hum Mutat*, 31, 477-483.
- Pavlidis, P. and Noble, W.S. (2003) Matrix2png: a utility for visualizing matrix data, *Bioinformatics*, 19, 295-296.
- Pruitt, K.D. et al. (2007) NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins, *Nucleic Acids Res*, 35, D61-65.
- Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features, *Bioinformatics*, 26, 841-842.