Zbrowse: an interactive GWAS results browser

Greg R Ziegler, Ryan H Hartsock, Ivan Baxter

The growing number of genotyped populations, the advent of high-throughput phenotyping techniques and the development of GWAS analysis software has rapidly accelerated the number of GWAS experimental results. Candidate gene discovery from these results files is often tedious, involving many manual steps searching for genes in windows around a significant SNP. This problem rapidly becomes more complex when an analyst wishes to compare multiple GWAS studies for pleiotropic or environment specific effects. To this end, we have developed a fast and intuitive interactive browser for the viewing of GWAS results with a focus on an ability to compare results across multiple traits or experiments. The software can easily be run on a desktop computer with software that bioinformaticians are likely already familiar with. Additionally, the software can be hosted or embedded on a server for easy access by anyone with a modern web browser.

Greg R Ziegler
greg.ziegler@ars.usda.gov
United States Department of Agriculture – Agricultural Research Service
St. Louis, MO

Ryan H Hartsock
rhartsock423@gmail.com
Donald Danforth Plant Science Center
St. Louis, MO

Ivan Baxter
ivan.baxter
United States Department of Agriculture – Agricultural Research Service
St. Louis, MO

Corresponding Author:
Ivan Baxter
975 North Warson Road
Saint Louis, MO 63132
Phone: 314.587.1438
Fax: 314.587.1538
ivan.baxter@ars.usda.gov

Title: Zbrowse: an interactive GWAS results browser

Abstract

The growing number of genotyped populations, the advent of high-throughput phenotyping techniques and the development of GWAS analysis software has rapidly accelerated the number of GWAS experiments. Candidate gene discovery from these datasets is often tedious, involving many manual steps searching for genes in windows around a significant SNP. This problem rapidly becomes more complex when trying to compare multiple GWAS studies to identify pleiotropic or treatment/environment specific effects. To address this problem, we have developed a fast and intuitive interactive browser for the viewing of GWAS results with a focus on an ability to compare results across multiple traits or experiments. The software can easily be run on a desktop computer

with open access software. Additionally, the software can be hosted or embedded on a server for easy access by anyone with a modern web browser.

Introduction

The recent development of high-throughput phenotyping techniques coupled with the ability to genotype large populations of diverse individuals has revolutionized the way that forward genetics research is performed. Tools have rapidly become available to perform genome-wide association studies (GWAS) in a variety of species (Kang et al., 2010; Segura et al., 2012; Lipka et al., 2012) that can map traits to the genome with high enough resolution to quickly provide a tractable list of potential causal genes.

One of the first steps an analyst will take is determining what gene or genes fall under the SNP peaks that can be seen on the Manhattan plot. Unfortunately, these plots are generally not interactive. Identifying the peaks of interest usually involves sifting through the results table for the range of coordinates under the peak of interest and then using those coordinates to filter a large gene annotation file. The extra steps involved in exploring the data in this way makes it more likely that interesting associations may be missed either due to 1) mistakes made in attempting to mine the large results files or 2) the dataset not being mined deeply enough due to the difficulty of looking for genes under less significant peaks. Additionally, this method quickly becomes tedious when analyzing multiple phenotypes or relatively complex traits.

Some web applications provide tools for viewing Manhattan plots (Table 1), but they are all either specific to a single species or don't allow interactive results browsing. These resources also do not allow for easy viewing and comparison of GWAS results across phenotypes and studies, a situation that frequently arises with structured populations.

Goals

We approached the construction of a new GWAS browser with the goal of giving the users the following tools, all of which were focused on versatility and adaptability:

1. *Ability to plot multiple traits in the same panel.* We wanted to enable users to find genotype-environment (GxE) interactions (e.g., those instances where an environmental condition causes a phenotypic effect, but only for individuals with a given allele) and loci with pleiotropic effects (the same loci affecting multiple phenotypes).
2. *Ability to rapidly move between scales (thousands of bps to billions).*

3. *Ability to find overlaps or commonalities among datasets.*
4. *Ability to interact directly with the plots*. We wanted the ability to look at the annotations of genes inline easily and link to additional information.
5. *Ability to plot both SNPs and genetic intervals*. We wanted users to be able to combine the results of quantitative trait locus mapping techniques with GWAS results.
6. *Ability to download plots and gene lists.*
7. Finally, we wanted all of this information and functionality to be available in one browser window using tools that are common and freely available on most personal computers.

Here, we present an interactive GWAS results viewer that is an extension of the classic GWAS Manhattan plot. It allows for the rapid comparison of GWAS results from multiple phenotyping experiments and the rapid viewing and analysis of genes under peak SNPs. *Arabidopsis thaliana*, maize, soybean, and sorghum are bundled with the software but we provide instructions and tools to easily add support for other organisms, including those with sex chromosomes. As a practical application of the browser's usage, we will demonstrate it using results from a recently completed sorghum GWAS experiment in which elemental profiling phenotypes were measured in accessions grown across three separate locations (Shakoor et al. 2015).

User Interface

The ZBrowse GWAS results viewer is an interactive application that runs on a local machine using R and is rendered in any modern web browser. Because the browser runs on the user's local machine, the data can remain private. Though the focus of the first version is a local installation, the browser display allows for easy sharing of the application. The browser is designed to be a streamlined environment that provides fast access to visualization tools for GWAS results. ZBrowse utilizes a tab-based navigation format to make accessing different aspects of the browser fast, efficient, and intuitive. There is also a sidebar panel on the left of the page that updates with a set of options specific to the tab being displayed.

The first tab in the list, and the landing page when the application is first loaded, is the Manage tab (Figure 1). This tab allows a new GWAS dataset to be uploaded into the application or a pre-loaded dataset from a dropdown menu can be selected. Data can be uploaded in a flat file delimited with either commas or tabs or an RData object. These flexible file formats allow any type of data to be loaded into the browser.

In Figure 1, we have loaded the results from the sorghum ionomics experiment and selected the appropriate columns to be used for plotting the results. The results file was generated by taking the most significant SNP hits from each of the 80 GWAS experiments performed (20 phenotypes measured in 3 locations and an experiment combining the location data). We added a column describing which experiment (e.g. the three locations) and which phenotype each SNP was found in.

Once uploaded, a preview of the first ten rows of the dataset will appear in the main panel. Below this table is a series of selection boxes that allow the user to specify which columns in the file to use for plotting. This selection method removes the complexity of requiring the input file to either have columns with specific names or columns in a specific order.

The user needs to select a chromosome and base pair for determining the location of each SNP in the genome. To plot base pair intervals, there is a separate checkmark box. Checking this box opens a second set of selection boxes allowing the user to select columns defining a start base pair and a stop base pair for each interval, as well as a separate y-axis column (allowing the traits to be plotted on different scales). Chromosomes expected for each organism are defined in a separate file (See Adding Organisms section). The browser supports both numeric and alphabetical chromosome names. So scaffold names (i.e. scaffold_1, scaffold_2), sex chromosomes (i.e. X, Y), or chromosome arm designations (i.e. 1S, 1L) are acceptable.

If the uploaded dataset is data from only one GWAS trait, there is a checkbox to include all data as one trait. Otherwise, the user can select one or more columns to group the data by when plotting. For example, a researcher might be interested in comparing GWAS results from multiple experiments, or in comparing results from multiple traits measured in the same experiment, or both. The user can designate as many columns as desired to subset the data that will be used to create tracks designated by different colors on the plots. Color coding in the plots for each track will match between the intervals and the single points. In our sorghum dataset, we are interested in exploring relationships between both the phenotypes measured and the effect environment may have had on our GWAS results. So we have selected the two columns that we added describing what location and phenotype each SNP result is from.

Finally, the user needs to select the y-axis column with the significance value against which to plot each SNP. Usually, this is the negative logarithm of the P-value, but can also be the number of bootstrap models that include this SNP (RMIP, Valdar et al., 2009) or any other measure of trait significance, such as effect size. If a user hasn't already created a column with the negative logarithm of the P-value, the P-value column can be selected and there is a checkbox to have the software automatically perform the calculation. The final parameter allows for user selectable values for the Y-axis scale. By default, the software will automatically scale the y-axis based on the range of the selected data. The browser will only display 5000 points total (see Limitations section). If there are more than 5000 points in the subset of tracks being plotted, the browser will use the y-axis column to rank the SNPs and take only the top 5000.

After the user has selected the appropriate parameters, clicking the submit button will trigger a tab change to the Whole Genome View visualization tab (Figure 2). Conveniently, once submitted, the software will remember the selected settings for this dataset on future visits and automatically populate the fields with the previously selected parameters. The plot on this tab is formatted as a standard genome-wide Manhattan plot. The x-axis is ordered by chromosome and base pair within each chromosome. The background of the plot has alternating blue/white shading for the even and odd chromosomes to highlight chromosome breaks. The panel on the left contains a box for each trait column selected in the Manage tab. In the case of our Sorghum ionomics dataset, there is a box where we can select which combination of the 20 phenotypes we would like to plot and which experimental location we would like to plot. In Figure 2, we have selected all four locations and we are plotting the significant associations found for cadmium, cobalt, and zinc. There is also an option for showing only overlapping SNPs with the ability to adjust both the overlap size around each point and the minimum number of overlaps.

When the user scrolls over points in the plot, a tooltip will display that shows information about the trait that SNP is associated with, the Y-axis value, and the exact chromosome and base pair for the SNP (Figure 2). If the tooltip gets in the way of the viewing or selecting of points, clicking the plot will temporarily hide the tooltip box. Clicking any point in the Whole Genome View will change tabs to the Chromosome View tab with a focus on the clicked point (Figure 3). In our example, we clicked on the peak SNP for cadmium in the Mexico experiment. This tab contains two plots: one is a chromosome-wide view displaying the data from the chromosome clicked in the genome-wide view, the other plot is an annotation plot of the region around the clicked base pair. A blue band in the chromosome-wide plot highlights the region being displayed in the annotation plot. In one click, we have gone from viewing the entire 700 million base pair genome, to a plot displaying the region 250 thousand base pairs around our point of interest. The plot contains a variety of interactive features. In addition to selecting traits to view in the sidebar panel, traits can be quickly hidden by clicking their entry in the legend. When many points are plotted on the same graph, overplotting can make it difficult to discern points clustered around the same peak. To alleviate this, the plot can be easily zoomed by clicking and dragging a zoom box anywhere in the plot. This makes it much easier to see the relationship between tightly grouped points. The displayed chromosome can be changed without returning to the Whole Genome View tab using the drop-down menu in the sidebar panel. Points can be clicked to redraw the annotation plot around new points of interest.

The annotation plot is a variable width plot that defaults to showing the region 250,000 base pairs on either side of the point of interest. The width of this region can be adjusted between 1,000 and 500,000 base pairs using the slider on the sidebar panel. The bottom of this plot has a track that shows the position of coding sequences around the SNP of interest. The tooltip for genes in this track displays information about the gene location, strand, and function, if known. For maize, arabidopsis and soybean, clicking on the gene will open a new browser tab that links to the gene description page specific to the organism being viewed. Arabidopsis links to The Arabidopsis Information Resource(TAIR) (Lamesch et al., 2011), soybean links to Soybase (Grant et al., 2010), and maize links to the Maize Genetics and Genomics Database (MaizeGDB) (Lawrence et al., 2004). In addition, clicking genes in organisms added from Phytozome (Goodstein et al., 2012) via the add organism application described below opens the Phytozome description page for that gene. ZBrowse can be easily modified to link out to other species-specific databases that can accept a query string in the URL.

Our cadmium example in Figure 3 shows how quickly we can find potential leads for candidate genes. Browsing the gene track a gene nearly directly under the peak cadmium SNP is annotated as being similar to a cadmium/zinc-transporting ATPase, clicking this gene opens a new browser tab displaying the phytozome gene description page.

In addition to the visual browser, annotation data can be explored in tabular form in the Annotation Table tab (Figure 4). This table provides an interactive table of the genes found in the window around the selected point. The table is sortable and searchable and can also be exported as a comma-separated file. A similar table viewer is available in the Data Table tab for analysis of the selected GWAS dataset.

Adding Organisms

Currently, maize, soybean, arabidopsis and sorghum are downloaded with the browser source package. We have developed an application to quickly add organisms to the browser from annotations downloaded from the Plant Genomics Portal (Phytozome) to the local installation of ZBrowse. Additionally, we will be formatting requested and popular organisms and releasing the files on GitHub. These will be easy to download and incorporate into your existing browser installation.

Adding a new organism manually requires two additional files to be created and placed into the ZBrowse installation directory. One is a flat text file with three lines. The first line tells the browser what the display name for the organism is. The second line tells the browser the names and size of each genome feature (i.e. chromosomes, scaffold, etc.) and the third line is the path to a csv file containing the annotation information. The annotation file needs to have the following columns: name, chromosome, transcript_start, transcript_end, strand, ID, defLine, bestArabhitDefline and bestRiceHitDefline.

Technical Foundation

The GWAS browser is written in the R programming language using packages that provide wrappers around popular javascript web applications including shiny (RStudio Inc., 2013) and rCharts (Vaidyanathan, 2013). Because of this, the browser can be run locally with only R and any modern web browser. Internal data processing makes use of the plyr package (Wickham, 2011). The javascript plots are drawn using Highcharts (highcharts.com) and are available for use under the Creative Commons Attribution-NonCommercial 3.0 License. Tables are generated using the javascript library Datatables (datatables.net) and xtable (Dahl, 2013). All of the tools and software used are either free or open source. The use of R to build the web application makes it more easily accessible to bioinformaticians to extend than if it was written in pure javascript. Many GWAS programs are written in R (Kang et al., 2008; Segura et al., 2012; Lipka et al., 2012). So, many scientists performing GWAS will already have some familiarity with R constructs, even if they are not computational biologists. This familiarity will hopefully make it easier for the community who is using the browser to extend it and modify it for their purposes.

Limitations

The browser takes a fundamentally different approach from current state of the art browsers. It is focused on the ability to quickly plot a variety of GWAS experiments on a single Manhattan plot. A caveat to this ability, however, is that it cannot plot every SNP in a genotype dataset. Due to memory, time, and plotting constraints the current browser is limited to approximately 5000 data points per trait, which is significantly less than most genotype datasets. Of course, only the most strongly associated SNPs are typically of interest, so this problem can be easily mitigated by trimming the input file to contain only significant associations (e.g., $p<0.05$). Currently, the browser will automatically trim the number of points being plotted to only display the top 5000 points based on the y-axis column. Future improvements to the browser could support the plotting of more information by binning points when zoomed out to a point where over plotting is an issue and only loading individual data points asynchronously when the zoom level is sufficient to see individual points.

The generality of the browser allows for it to be used with any SNP dataset. Only chromosome number and base pair information needs to be provided for each SNP. However, this means that specific information about the genotype dataset being used, such as minor allele frequency or linkage disequilibrium information, cannot be displayed on the plot. Of course, the flexibility of the browser would make it easy to build personalized solutions that could display additional information for specific SNP datasets and additional tracks could be added to display linkage disequilibrium decay around the selected SNP.

One obvious extension of the browser that would address many of the limitations listed above would be to connect it to a database designed to quickly and efficiently handle all of the data that goes into a GWAS experiment. Database support would allow custom subsetting of entire GWAS datasets and if the GWAS genotype files are available, then summary data about each particular SNP could also be displayed. This would allow the browser to be incorporated into a much larger ecosystem that could take a GWAS experiment from phenotypic dataset, through running a GWAS experiment, to final analysis and visualization.

While the limitations identified above may constrain the use of the browser for certain applications, there are a number of use cases that are enabled by its current functionality. Using open source tools and GitHub for the code distribution, the browser functionalities can be enhanced by the authors or by other members of the user community.

Available resources:

Download at:

http://www.baxterlab.org/#!/cqi0

Code is also available on github at

https://github.com/baxterlabZbrowse/ZBrowse

manual can be found here:

http://media.wix.com/ugd/52737a_2a65d0deb3bd4da2b5c0190c0de343ca.pdf

For support email:

baxterlabZbrowse@danforthcenter.org

References

Beck T, Hastings RK, Gollapudi S, Free RC, Brookes AJ. 2014. GWAS Central: a comprehensive resource for the comparison and interrogation of genome-wide association studies. *European journal of human genetics: EJHG* 22:949–952.

Cuellar-Partida G, Renteria ME, MacGregor S. 2015. LocusTrack: Integrated visualization of GWAS results and genomic annotation. *Source Code for Biology and Medicine* 10:1.

Dahl DB. 2013. *xtable: Export tables to LaTeX or HTML*.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS. 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research* 40:D1178–D1186.

Grant D, Nelson RT, Cannon SB, Shoemaker RC. 2010. SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Research* 38:D843–D846.

Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E. 2008. Efficient Control of Population Structure in Model Organism Association Mapping. *Genetics* 178:1709–1723.

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S, Freimer NB, Sabatti C, Eskin E. 2010. Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics* 42:348–354.

Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, Karthikeyan AS, Lee CH, Nelson WD, Ploetz L, Singh S, Wensel A, Huala E. 2011. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research* 40:D1202–D1210.

Lawrence CJ, Dong Q, Polacco ML, Seigfried TE, Brendel V. 2004. MaizeGDB, the community database for maize genetics and genomics. *Nucleic Acids Research* 32:D393–D397.

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z. 2012. GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28:2397–2399.

Li MJ, Sham PC, Wang J. 2012. Genetic variant representation, annotation and prioritization in the post-GWAS era. *Cell Research* 22:1505–1508.

Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. 2010. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics (Oxford, England)* 26:2336–2337.

RStudio Inc. 2013. *shiny: Web Application Framework for R*.

Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, Nordborg M. 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nature Genetics* 44:825–830.

Seren Ü, Vilhjálmsson BJ, Horton MW, Meng D, Forai P, Huang YS, Long Q, Segura V, Nordborg M. 2012. GWAPP: A Web Application for Genome-Wide Association Mapping in Arabidopsis. *The Plant Cell Online*:tpc.112.108068.

Shakoor N, Ziegler G, Dilkes BP, Brenton Z, Boyles R, Connolly EL, Kresovich S, Baxter I. 2015. Integration of experiments across diverse environments identifies the genetic determinants of variation in Sorghum bicolor seed element composition. bioRxiv doi:10.1101/019083

Skinner ME, Uzilov AV, Stein LD, Mungall CJ, Holmes IH. 2009. JBrowse: A next-generation genome browser. *Genome Research* 19:1630–1638.

Vaidyanathan R. 2013. *rCharts: Interactive Charts using Polycharts.js*.

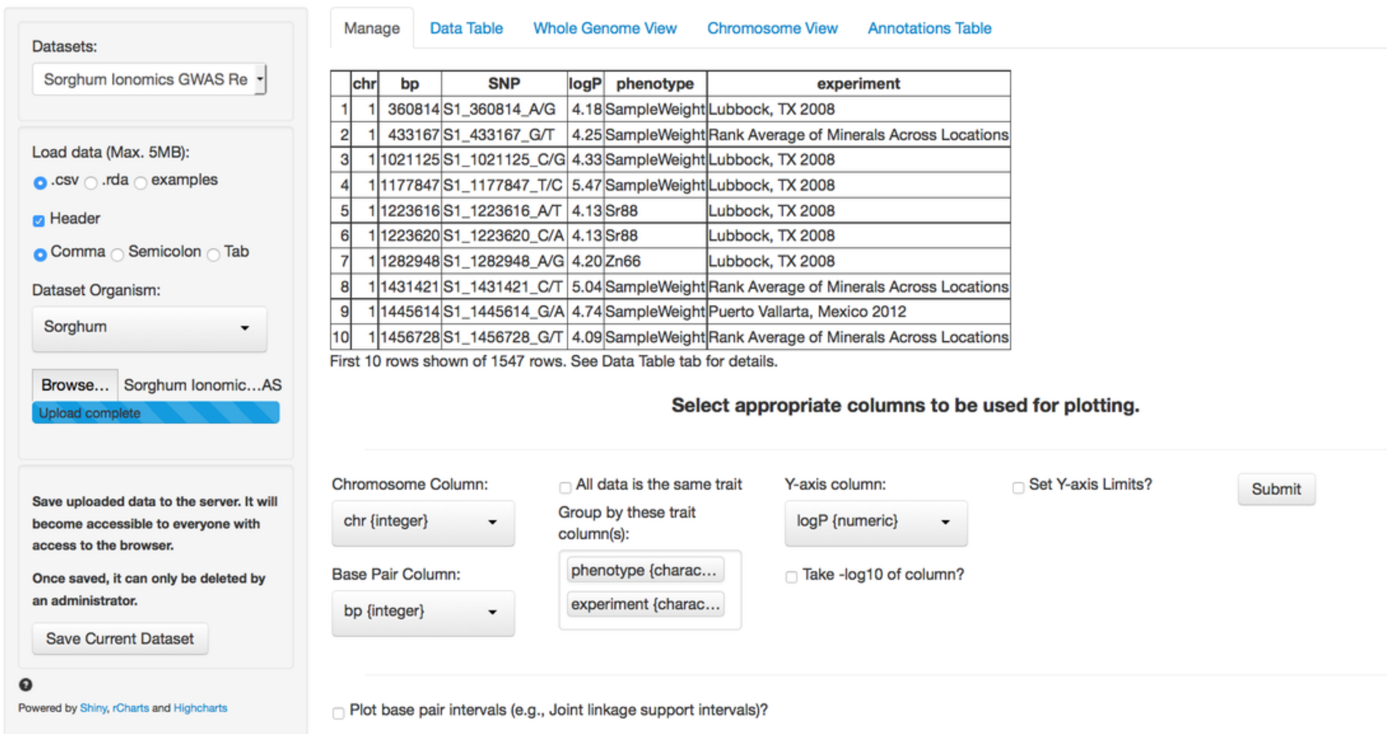
Valdar W, Holmes CC, Mott R, Flint J. 2009. Mapping in Structured Populations by Resample Model Averaging. *Genetics* 182:1263–1277.

Wickham H. 2011. The Split-Apply-Combine Strategy for Data Analysis. *Journal of Statistical Software* 40:1–29.

1

Landing page for ZBrowse.

A sorghum GWAS dataset has been uploaded and the appropriate columns for plotting the data have been selected from the selection boxes.

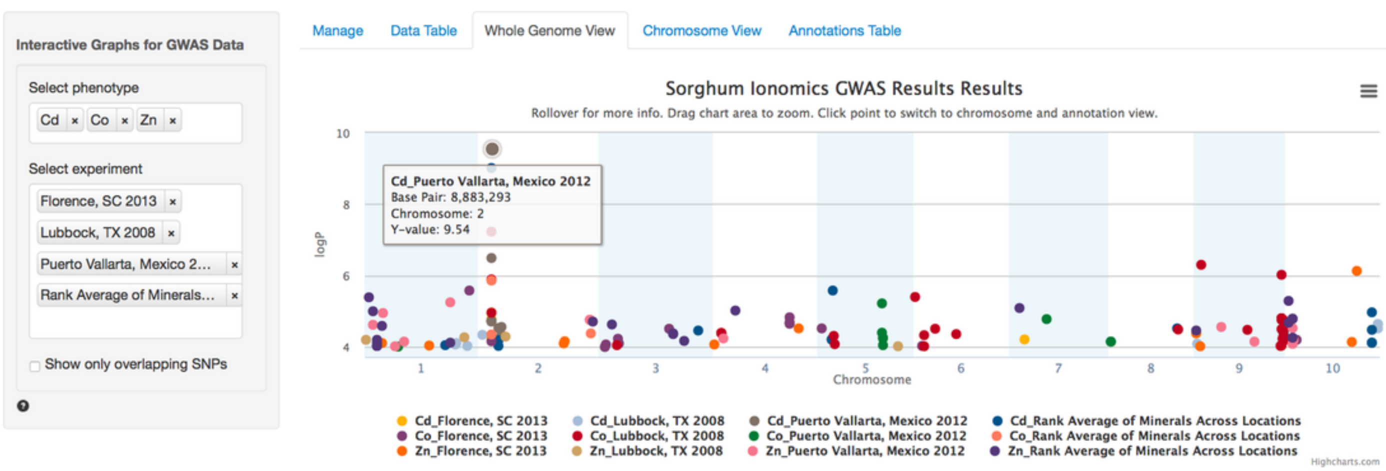


	chr	bp	SNP	logP	phenotype	experiment
1	1	360814	S1_360814_A/G	4.18	SampleWeight	Lubbock, TX 2008
2	1	433167	S1_433167_G/T	4.25	SampleWeight	Rank Average of Minerals Across Locations
3	1	1021125	S1_1021125_C/G	4.33	SampleWeight	Lubbock, TX 2008
4	1	1177847	S1_1177847_T/C	5.47	SampleWeight	Lubbock, TX 2008
5	1	1223616	S1_1223616_A/T	4.13	Sr88	Lubbock, TX 2008
6	1	1223620	S1_1223620_C/A	4.13	Sr88	Lubbock, TX 2008
7	1	1282948	S1_1282948_A/G	4.20	Zn66	Lubbock, TX 2008
8	1	1431421	S1_1431421_C/T	5.04	SampleWeight	Rank Average of Minerals Across Locations
9	1	1445614	S1_1445614_G/A	4.74	SampleWeight	Puerto Vallarta, Mexico 2012
10	1	1456728	S1_1456728_G/T	4.09	SampleWeight	Rank Average of Minerals Across Locations

2

Genome wide view of ZBrowse manhattan plot.

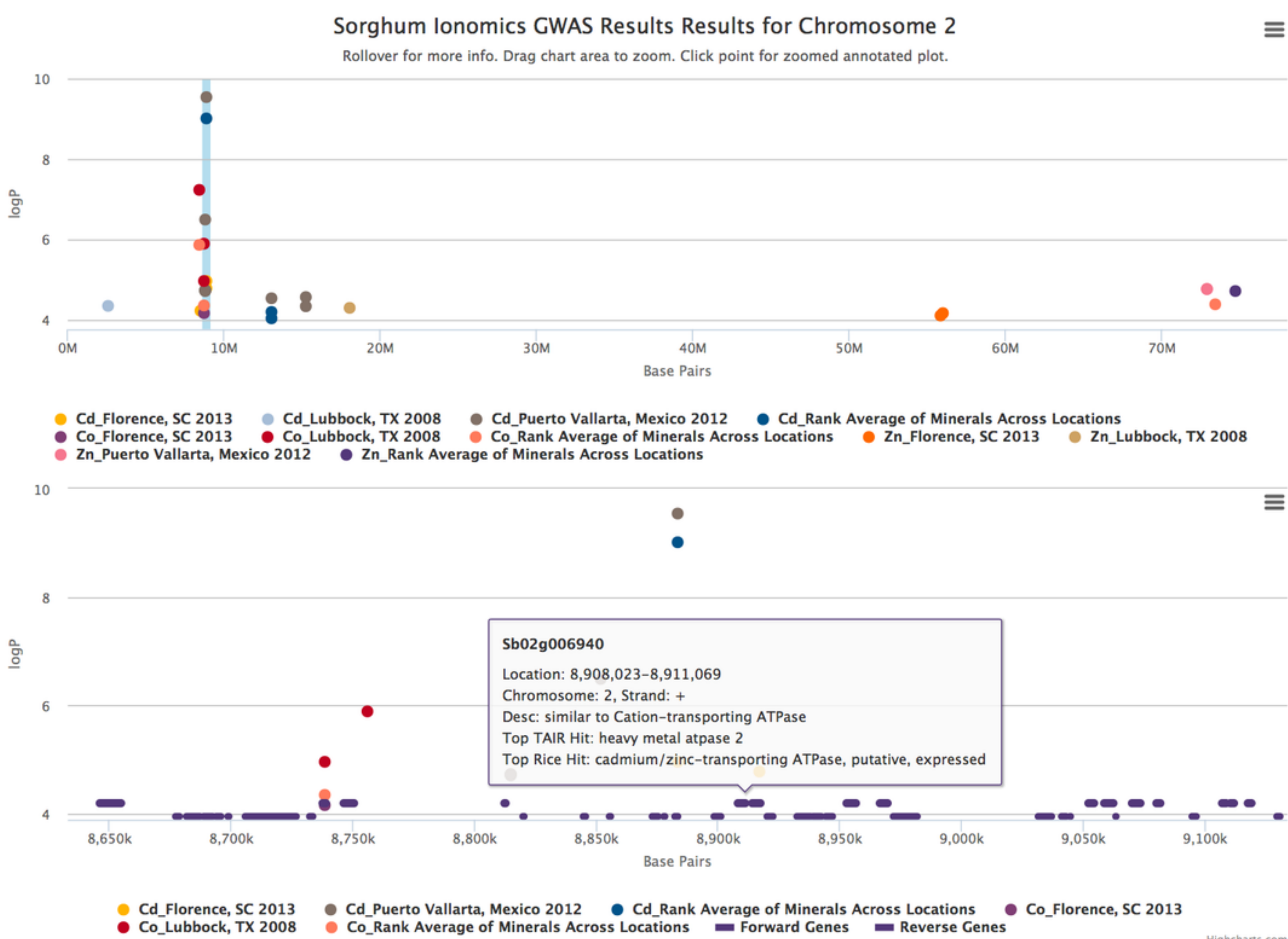
The plot is displaying a comparison of GWAS results from three phenotypes measured in three separate locations and one aggregate experiment (12 GWAS experiments total on one graph). The legend at the bottom of the figure displays the color of points that correspond to the combination of traits and locations selected in the sidebar on the left hand side of the figure. Clicking the points in the legend allows a user to easily show or hide points from that trait. The title of the plot is automatically generated from the filename of the dataset provided by the user. This makes it easy to determine which GWAS experiment is being plotted. The tooltip popup is displayed by hovering the mouse over points in the plot.



3

Chromosome and annotation level manhattan plots in ZBrowse.

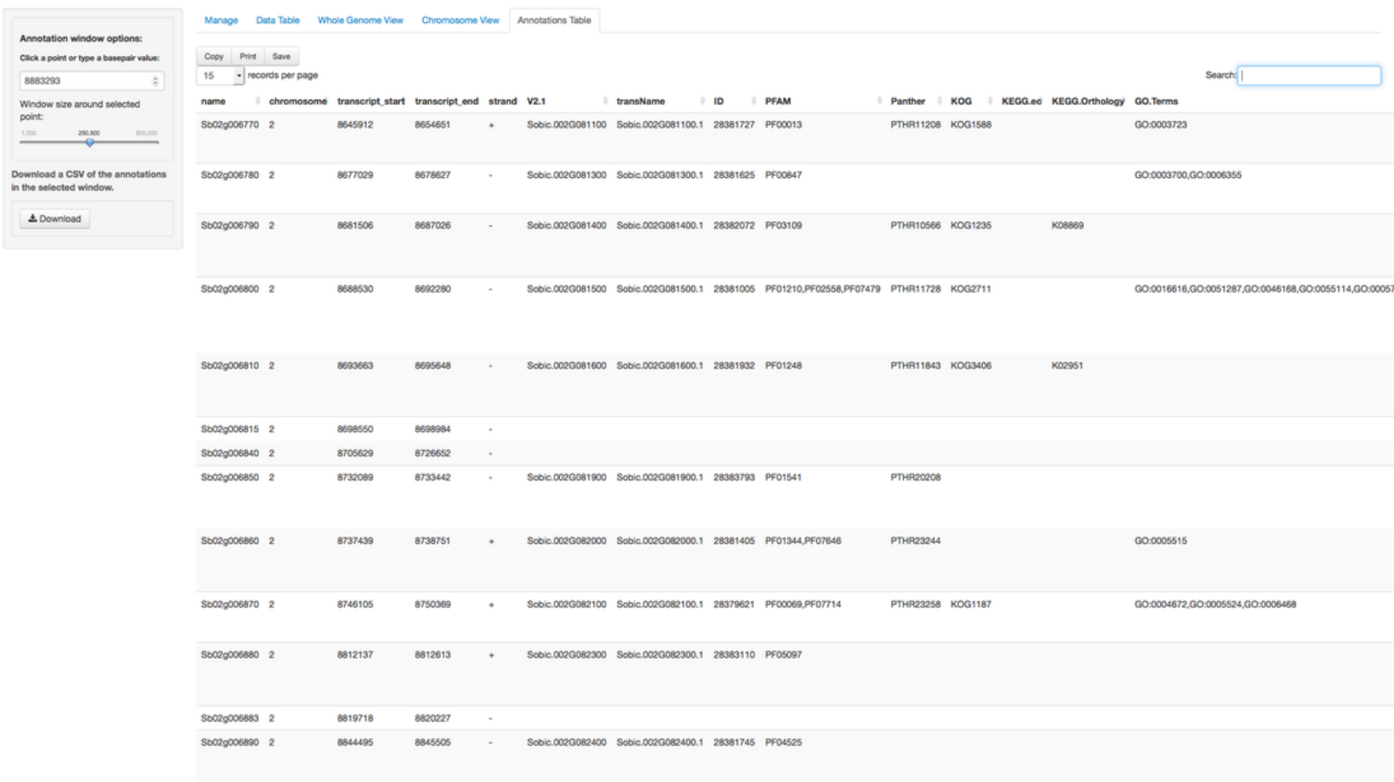
This plot was reached by clicking the peak SNP for cadmium in the Puerto Vallarta, Mexico experiment displayed in Figure 2. The blue vertical bar in the upper chromosome level plot indicates the zoomed region in the annotation level plot below. Two tracks at the bottom of the annotation level plot indicate locations of genes on the forward and reverse strands. Scrolling over these tracks displays a tooltip with a description of the gene and clicking genes in the track opens a separate browser window displaying information about the gene from an external database. The displayed gene, as heavy metal transporter, is a likely candidate for effecting cadmium accumulation in sorghum germplasm.



4

Annotation Table tab in ZBrowse.

Tabular form of the genes found around a SNP. This table is searchable in the browser and can be exported as a csv formatted file.



name	chromosome	transcript_start	transcript_end	strand	V2.1	transName	ID	PFAM	Panther	KOG	KEGG.ec	KEGG.Orthology	GO.Terms
Sb02g006770	2	8645912	8654651	+	Sobic.002G081100	Sobic.002G081100.1	28381727	PF00013	PTHR11208	KOG1588			GO:0003723
Sb02g006780	2	8677029	8678627	-	Sobic.002G081300	Sobic.002G081300.1	28381625	PF00847					GO:0003700,GO:0006355
Sb02g006790	2	8681506	8687026	-	Sobic.002G081400	Sobic.002G081400.1	28382072	PF03109	PTHR10566	KOG1235		K08869	
Sb02g006800	2	8688530	8692280	-	Sobic.002G081500	Sobic.002G081500.1	28381005	PF01210,PF02558,PF07479	PTHR11728	KOG2711			GO:0016616,GO:0051287,GO:0046168,GO:0055114,GO:000573
Sb02g006810	2	8693663	8695648	-	Sobic.002G081600	Sobic.002G081600.1	28381932	PF01248	PTHR11843	KOG3406		K02951	
Sb02g006815	2	8698550	8698984	-									
Sb02g006840	2	8705629	8726652	-									
Sb02g006850	2	8732089	8733442	-	Sobic.002G081900	Sobic.002G081900.1	28383793	PF01541	PTHR20208				
Sb02g006860	2	8737439	8738751	+	Sobic.002G082000	Sobic.002G082000.1	28381405	PF01344,PF07646	PTHR23244				GO:0005515
Sb02g006870	2	8746105	8750369	+	Sobic.002G082100	Sobic.002G082100.1	28379621	PF00069,PF07714	PTHR23258	KOG1187			GO:0004672,GO:0005524,GO:0006468
Sb02g006880	2	8812137	8812613	+	Sobic.002G082300	Sobic.002G082300.1	28383110	PF05097					
Sb02g006883	2	8819718	8820227	-									
Sb02g006890	2	8844495	8845505	-	Sobic.002G082400	Sobic.002G082400.1	28381745	PF04525					

Table 1(on next page)

Comparison of ZBrowse with other available GWAS visualization software packages.

LocusZoom (Pruim et al., 2010), LocusTrack (Cuellar-Partida, Renteria & MacGregor, 2015), GWAS Central (Beck et al., 2014), GWAPP (Seren et al., 2012), GWASrap (Li, Sham & Wang, 2012), JBrowse (Skinner et al., 2009)

Browser	Language programmed in	Is it run on local, server or web	Is the plot interactive	Display SNPs and intervals	Organisms Supported	Displays multiple GWAS experiments
LocusZoom	R, Python wrapper, SQLite table	Local or web	No	No	Human	No
LocusTrack	R	Local or web	No	No	Human	No
GWAS Central	Unclear	Web	Partial	No	Human	No
GWAPP	Python, HTML5, Javascript	Web	Yes	No	Arabidopsis	No
GWASrap	Unclear	Web	Yes	No	Human	No
JBrowse	Perl/Javascript	Local, server or web	Yes	Yes	Any	Not easily on the same track
ZBrowse	R	Local, server with some modification	Yes	Yes	Any	Yes

2