# **Practical II – Expression browsers, KnetMiner, and using TILLING resources**

The aim of this session is to learn how to use the eFP and expVIP expression browsers, KnetMiner, as well as making use of TILLING resources for wheat.

## **Background information for the exercise**

You have mapped a QTL for grain length to a distinct genetic interval on chromosome 6A. You now want to understand the expression patterns of the candidate genes in the interval, learn more about what is known in wheat and other species about them, as well as to develop functional knockout mutants using TILLING.

## **Practical exercise 1 – Expression browsers**

**Expression analysis**

**eFP browser**

* Go to the eFP browser at <http://bar.utoronto.ca/efp_wheat/cgi-bin/efpWeb.cgi>. Note that the eFP browser only accepts gene models using the RefSeqv1.0 annotation, e.g. **TraesCS2B01G000100**.
* Let us look at the expression pattern of one of the genes in the QTL interval. Enter “TraesCS6A01G313800” as the primary gene ID and press “Go” to investigate the gene in absolute mode. Hover over the various tissues to inspect the expression values and variances. Which tissue has the highest expression? Which the lowest?
* You can download all values quickly as a table. To do that, scroll to the bottom of the page and click the “Table of Expression Values” button. It will open a pop-up window showing the expression values plus standard deviation. This can help you find high and low expression values easily.
* Look at the expression of the gene using a different dataset. Click on the box labelled “Data Source” and select the “Wheat Embryogenesis” dataset. At which stage during embryogenesis does the gene’s expression reach its maximum?
* This expression browser is well and good but let us try something that offers some more flexibility in how the data is presented.

**expVIP browser**

* Navigate to the expVIP website for wheat (<http://www.wheat-expression.com>). Note that expVIP supports CSS, TGAC, as well as RefSeqv1.0 and RefSeqv1.1 gene names.
* Type in the name of one of the genes from the QTL interval, TraesCS6A02G313800, and click the “Search” button.

* expVIP is a very flexible expression browser that uses “metadata” categories to combine or divide the data according to your selections. You can combine or expand datapoints by “closing” or “opening” the metadata categories on the left-hand side using the + and – signs. By default, expVIP opens up the “high level tissue”, “high level age”, “high level stress-disease”, and “high level variety” metadata categories.

You can further sort the order in which the expression values are shown by clicking on the rows of coloured dots on the left.

For now, open the “Age” and “Tissue” categories and close all other ones, then sort by “Tissue”.

* Inspect the expression values of the gene in the different tissue/age combinations. Which tissue shows the highest expression at what age? Which tissue shows the lowest expression?
* Click the box labelled “Homoeologues” to view expression of your genes homoeologs. Compare the expression patterns of the three homoeologs. Are the genes expressed in a similar fashion? Are there any tissue/age combinations where expression is different among the homoeologs?
* Copy the genes from the QTL interval (see below) into the “Multiple Genes” box and inspect the output. Which of these genes are highly expressed in grains?

TraesCS6A02G312300 TraesCS6A02G313500

TraesCS6A02G312400 TraesCS6A02G313600

TraesCS6A02G312500 TraesCS6A02G313700

TraesCS6A02G312600 TraesCS6A02G313800

TraesCS6A02G312700 TraesCS6A02G313900

TraesCS6A02G312800 TraesCS6A02G314000

TraesCS6A02G312900 TraesCS6A02G314100

TraesCS6A02G313000 TraesCS6A02G314200

TraesCS6A02G313200 TraesCS6A02G314300

TraesCS6A02G313300 TraesCS6A02G314400

TraesCS6A02G313400

## **Practical exercise 2 – KnetMiner**

* Navigate to the KnetMiner website for wheat (<https://knetminer.com/Triticum_aestivum/>) and enter all the gene names from the QTL interval (see last section of previous exercise) into the search box.
* Inspect the connections, what are some putative functions of these genes?
* Filtered the result by keywords. Try “grain length”, “grain size”
* Which gene do you think is the most likely candidate based on the identified connections and why? What is some evidence that supports the putative function of these genes? Are you convinced?

## **Practical exercise 3 – TILLING resources**

* Navigate to the *Ensembl* Plants website (<http://plants.ensembl.org/index.html>) and enter “TraesCS6A02G313800” into the search box.
* Click on the “Gene ID” link to get to the genic information for “TraesCS6A02G313800”.
* On the left-hand side tab, click on “Variant image”. This shows you all mutations (including EMS) covering your gene-of-interest. However, it is difficult to find individual mutations.
* On the left-hand side tab, click on “Variant table”. This is the same data as seen under “Variant image”, but in an ordered tabular format.
* Filter the table by selecting only “EMS-induced mutation” under “Source”. Further filter for mutations in tetraploid “Kronos” wheat by typing “Kronos” into the search box just above the table.
* Select two TILLING mutants that you think could lead to a knock-out of the gene’s function. Save the variant names (e.g. Kronos4262.chr6B.291763807) in a text file. Click on each variant to get more information about them. Are they homo- or heterozygous? Is there already a marker designed for the mutation?
* Select two TILLING mutants for the B-genome homoeolog “TraesCS6B02G343900”.
* Go to BioMart and select “Ensembl Plants Variation 56” as the database and “Triticum aestivum Short Variants (SNPs and indels excluding flagged variants) (IWGSC)” as your dataset.
* Under “Filters” and then “GENERAL VARIANT FILTERS”, enter the names of your selected variants into the “Filter by Variant name (e.g. rs123, CM000001) [Max 500 advised]” box.

Under “Attributes”, select “Flanking Sequences” and tick the “Variant sequences” as well as “Upstream flank” and “Downstream flank” boxes. Select 100 bp for the up- and downstream flanks.

* Click “Results” and you will find information for each of your variant plus flanking sequences. Where the SNP is located, the sequence shows an underscore “\_”.
* Go to the PolyMarker website (<http://www.polymarker.info/>) and use the data from BioMart to design primers. Make sure to select “Tetraploid wheat, based on Chinese Spring RefSeqv1.0” as your reference.

Remember:

The necessary information for PolyMarker has the following format:

ID**,**chromosome-of-interest**,**flanking-sequence**[WT/Mut]**flanking-sequence

e.g.

Kronos4262.chr6B.291763807**,**6B**,**TTGTGTTGATATTAGAAAGCCAAATCATTTACTTTATCTTGTATACATTTTGTTACAGGAAGAACAGAAAGTCATTGAAGCACAGATGAGGATGCGCCAG**[C/T]**AAGCACTTCAAGACGAAGAGGATAAGATGAAAAGAAAACAGAGTAGGTGCTCTTCTAGCAGAACAATCGCTCCAACAACAGAAGTGGAGTATCGAGATAT