QTL Narrowing Tool (QNT) – User Manual
Alpha-release

The QNT Team: Dave Walton, Timothy Stearns, Annerose Berndt

Creation Date: June 30, 2009

Last Updated: September 21, 2010

Version: alpha

Table of Contents

A) Purpose and Significance 2

B) Overview of how QNT works 2

i) Input data by the user: 2

ii) The step-by-step procedures performed by QNT are: 2

iii) The results provided by QNT are presented in 3 tables: 3

C) Step-by-step guidance for the use of the QNT 4

i) Generate the input QTL list 4

1. Upload a user-defined QTL list. 4
2. Search the MGI database for QTL of interest. 4
3. Combine option 1. and 2. 5

ii) Modify the input QTL list table 5

iii) Microarray data input 5

iv) Narrow list of QTL 6

D) Additional functionality of the QNT for future releases 8

A) Purpose and Significance

The quantitative trait locus (QTL) Narrowing Tool (QNT) is a web-based application that integrates several bioinformatics tools with the goal to narrow single or multiple (overlapping) QTL regions to identify a list of potential candidate genes. This tool is intended to substitute the otherwise tedious manual QTL narrowing process and to standardize the methodology. The significant advantage of the QNT over manual QTL narrowing is its time- and labor-efficiency, which is achieved by:

- the utilization of QTL, sequence and expression database information in one tool;
- the automatic and gene-centered inter-species genome conversion.

B) Overview of how QNT works

i) Input data by the user:

1) The only body of data needed to use QNT is a list of QTL.

2) Additionally, the user has the option to incorporate microarray expression data into the QNT. Here the user has either the option to use already provided microarray data (we currently provide data from a strain survey for lung and liver gene expression) or user-defined microarray expression data through RMA file upload.

ii) The step-by-step procedures performed by QNT are:

1) Generation of the smallest common region between overlapping QTLs within species.

2) Haplotype analysis for smallest common region within mouse.

3) Inter-species genome conversions and identification of the smallest common region between-species.

4) Identification of all polymorphic SNPs within the regions.

5) Identification of all genes within the smallest common region between species.

6) Annotation of each polymorphic SNP with CGD SNPDB annotations and association with genes.

7) Association with probe set IDs using MGI (Mouse Genome Informatics).

8) Annotation of each gene with gene expression information.

iii) The results provided by QNT are presented in 3 tables:

1) A Region Table of the smallest common regions identified by the QTL narrowing process. This table lists the base pair range of each region, the QTL that were narrowed, the total number of SNPs within the region, the number of polymorphic SNPs and the identified genes. The base pair range provides a link to the UCSC Genome Browser for that range. The genes within each region are then the entry to the next table:

2) The Gene Table. This table lists each gene within a region with its symbol and name, a link to the individual MGI gene detail page and the number of polymorphic SNPs within the gene. If mircoarray data were used for the narrowing process the mean intensities of the high and low responding strains as well as the *P*-value of their difference are given. The polymorphic SNPs within each gene are then the entry to the next table:

3) The SNP Table. This table lists all the polymorphic SNPs within a gene with its base pair position, the high and low responder allele, the rs number (rs numbers are taken from the CGD SNP database), a secondary SNP ID and its source, and its annotation.

C) Step-by-step guidance for the use of the QNT

The web-based interface of the QNT is located at: http://jaxbhqtl1.jax.org:8080/QTLNarrowing/

To successfully use the QNT, please follow the following steps:

i) Generate the input QTL list

There are 3 option of how you can generate a QTL list:

1. **Upload a user-defined QTL list.**
    a. Prepare and save the QTL list file on your local wok station. To prepare a QTL list file follow these requirements for the custom QTL list file:
        i. The file needs to be txt-formatted (tab-delimited).
        ii. The file should not include a header line.
        iii. The file should include information in the first eight columns as follows:
            1. **QTL ID**: a specific ID to identify the QTL in comparison to other QTLs used in the same list: the ID is for the users purpose only and does not have to be the published QTL ID.
            2. **Phenotype**: the phenotype which was used to identify the QTL;
            3. **Species**: the species in which the QTL was identified;
            4. **High Resp Strain***: Strain name of the strain with the quantitatively higher phenotype; A list of valid strain names is available at http://jaxbhqtl1.jax.org:8080/QTLNarrowing/strains.txt
            5. **Low Resp Strain***: Strain name of the strain with the quantitatively lower phenotype; A list of valid strain names is available at http://jaxbhqtl1.jax.org:8080/QTLNarrowing/strains.txt
            6. **Chr**: chromosome (only numerical, for example: chromosome 1needs to be a "1" not a "chr1" or "Chr1");
            7. **QTL Start**: start position of the QTL or 95% confidence interval (in million base pairs (Mb));
            8. **QTL End**: end position of the QTL or 95% confidence interval (in million base pairs (Mb))
    b. To upload the QTL list file:
        i. Click the "Choose File" button.
        ii. Browse to the directory where you saved the QTL list file.
        iii. Select the QTL list file.
        iv. Click the "Upload" button.

2. **Search the MGI database for QTL of interest.**
    a. Click on the "Search MGI" button.
    b. Select the Phenotype(s) of interest.
    c. Upload the isolated QTL list by clicking on the "Upload QTL List" button.
    d. Click "Upload" button.

3. **Combine option 1. and 2.**
    a. Follow step 1. and 2. above.

ii) Modify the input QTL list table

Once the desired QTL list is uploaded into the QTL list table a variety of modification of this table can be undertaken to verify the correctness and modify the content if necessary.

The following modifications are possible:

1. Sort: By clicking on the drop-down menu on each column header you can sort in ascending or descending order.
2. Content: By clicking on the drop-down menu on each column header you can select "Columns". The list of all columns within the table will appear in a drop-down menu. By checking or un-checking the box next to a column you can decide which column is visible in the table. Un-checking does not mean deleting a column.
3. Strain names: The spelling of the strain names is crucial for the successful performance of the QNT. Incorrectly spelled strains will appear in red font. To correct the strain name please click on the individual strain cell and select the strain of interest from the drop-down menu. Please ensure that you scroll through your list of QTL completely to detect spelling errors. In case you miss a spelling error the narrowing process will not start and you will be asked to reevaluate the list of QTL.

iii) Microarray data input

The purpose of using the microarray expression data is to give the user an additional line of evidence for or against a potential candidate gene.

The following options are possible:

1. Select one of the available microarrays from the drop-down menu. Caution: The provided microarrays are exclusively for lung or liver tissue. If you study other tissues be cautious about interpreting the data!
    a. In both provided microarrays the average signal intensities for each probe set within each arrays were calculated by the RMA function provided within the Affymetrix package for R using a custom (Entrez Gene) CDF file (Dai et al., 2005). The RMA method incorporates convolution background correction, sketch-quantile normalization, and summarization based on a multi-array model fit robustly using the median polish algorithm. (Dai, M., et al. 2005. Evolving Gene/Transcript Definitions Significantly Alter the Interpretation of GeneChip Data. Nucleic Acid Research. 33(20), e175.)

b. Overview of the experiments performed for each microarray experiment:

Tissue	Age of mice	Gender	Diet	Strains available
Lung	11 weeks	Females	6% Chow	C57BL/6J, SJL/J, BALB/cJ, C3H/HeJ, 129S1/SvlmJ, A/J, CE/J, PWD/PhJ, MRL/MpJ, DBA/2J, Ln/J, KK/HLJ
Liver	8 weeks	Females and Males	6% Chow and High fat	129S1/SvImJ, A/J, C57BL/6J, BALB/cJ, C3H/HeJ, CAST/EiJ, DBA/2J, I/LnJ, MRL/MpJ, NZB/BINJ, PERA/EiJ, SM/J

2. Upload your own RMA file. You will have the option to upload your own RMA files and have them analyzed by the QNT in an extra step. The analysis will follow the steps used for the microarrays above: the average signal intensities for each probe set within each arrays will be calculated by the RMA function provided within the Affymetrix package for R using a custom (Entrez Gene) CDF file
    a. Select the "Upload RMA File" radio button.
    b. Click "Choose File".
    c. Browse your local directory to the location of the RMA file and select.
        i. Requirements for the RMA file are:
            1. The RMA file should be a text tab-delimited file.
            2. The first column will need to be the probe ID with the remaining columns containing the intensities per array (each array is an additional column).
            3. Ideally, the data should be normalized prior to loading, but you may choose to do this within the QNT.
    d. Click "Define Exp. Design" and prepare the design of your expression experiment.

3. You can also choose not to include any microarray expression data by selecting the "No Gene Expression Comparison" radio button.

iv) Narrow list of QTL

To initiate the tool, click on the "Narrow QTLs" button. If the tool started narrowing successfully a processing dialog box will appear, which will give the current state of the narrowing process. The dialog

box is modal, and will not disappear until the analysis is done. You are still able to use your browser, but selecting another tab, or opening another window but the specific tab the analysis is running in will be blocked until the narrowing is completed.

The time required to narrow a QTL list depends of the number of QTLs and the overlap between QTLs.

If the narrowing does not start a dialog box will appear letting you know the error. Please correct the error and click “Narrow QTLs” again.

You have also the option to upload a new QTL list after clearing the QTL list table by clicking the “Clear” button.

D) Additional functionality of the QNT for future releases

The user will realize that some of the functionality of the tool is not implemented for the current alpha-release. For later releases of the QNT we will add the following functions:

1. MGI search
2. Inter-species comparison
3. User-defined RMA file upload

We appreciate any additional suggestions to improve the efficiency and the functionality of the QNT.

Thank you,
The QNT Team