CVD: the intestinal crypt/villus in situ hybridization database

Line Olsen¹, Morten Hansen¹, Claus T. Ekstrøm², Jesper T. Troelsen¹ and Jørgen Olsen¹,*

¹Department of Medical Biochemistry and Genetics, Biochemistry Laboratory C and ²Department of Biostatistics, Institute of Public Health, University of Copenhagen, The Panum Institute, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark

Received on August 14, 2003; revised and accepted on October 19, 2003
Advance Access publication February 10, 2004

ABSTRACT

Summary: The intestinal crypt/villus in situ hybridization database (CVD) query interface is a web-based tool to search for genes with similar relative expression patterns along the crypt/villus axis of the mammalian intestine. The CVD is an online database holding information for relative gene expression patterns in the mammalian intestine and is based on the scoring of in situ hybridization experiments reported in the literature. CVD contains expression data for 88 different genes collected from 156 different in situ hybridization profiles. The web-based query interface allows execution of both single gene queries and pattern searches. The query results provide links to the most relevant public gene databases.

Availability: http://pc113.imbg.ku.dk/ps/

Contact: jolsen@imbg.ku.dk

Large numbers of gene expression patterns unfold during the lifetime of a multicellular organism and it was recently pointed out that the complexity of gene expression patterns increases with increasing morphological and behavioural complexity of the organism (Levine and Tjian, 2003). The description of gene expression patterns with respect to temporal and spatial development is therefore of pivotal importance in order to understand the biology of multicellular animals. In situ hybridization is a high-resolution gene expression analysis technique that allows the steady state concentration of a specific mRNA to be monitored at the cellular level in tissue sections [for a description of the technique, see Ausubel et al. (1995)]. The histological preparation for each organ under study is often standardized, which ensures that the same tissue structures are present irrespective of the gene being investigated. In principle, this should facilitate comparisons of relative gene expression levels for different genes based on data already presented in the scientific literature. Unfortunately, however, no method exists for performing queries in the literature for specific gene expression patterns obtained from in situ hybridization studies. Using the mammalian small intestinal epithelium as a model system, we suggest here a general approach for organizing in situ hybridization data from the literature.

The epithelium of the mammalian small intestine (for a review, see Madara and Trier, 1987) is divided into discrete and functionally different compartments (see the CV navigator at http://pc113.imbg.ku.dk/ps/). The crypts contain at their bottom a few stem cells that give rise to daughter cells migrating towards the crypt openings. During the migration the daughter cells proliferate, migrate and differentiate. At the openings of the crypts the epithelium continues onto finger-like projections: the villi. The epithelial cells positioned at the villus base are differentiated and express digestive enzymes. At the tip of the villi, the cells undergo apoptosis and are shed into the intestinal lumen. The migration of an epithelial cell from the bottom of the crypt to the tip of the villi takes ~48 h depending on the species, and the epithelium as such is therefore constantly renewed. The dynamic structure of the small intestinal epithelium has made it a popular model to study gene regulation during cellular differentiation (for a review, see Stappenbeck et al., 1998) and we recently demonstrated a general decrease in transcriptome complexity during differentiation of the mouse small intestinal epithelium using filter-based cDNA arrays (Tadjali et al., 2002).

It was the purpose of the present work to construct a database for gene expression patterns in the small intestinal epithelium based on the in situ hybridization data already available in the scientific literature and to provide a user-friendly web-based interface to query the database with emphasis on easy retrieval of genes with similar relative expression patterns along the small intestinal crypt/villus axis.

By searching PubMed release 200801, 145 references for papers reporting on in situ hybridization in the mammalian intestine were identified. The references represented 88 different genes and 156 figures with in situ hybridization images. A smaller fraction of the references dealt with the colonic epithelium and it was decided to include this.

*To whom correspondence should be addressed.
information in the database as well. To organize the data, the crypt/villus axis was divided into four segments: the bottom crypt segment corresponding to the stem cell compartment; the upper crypt region where the committed undifferentiated crypt cells reside; the lower villus region containing the differentiated absorptive epithelial cells; and the upper villus region where cell shedding takes place. A simple relative scoring scheme was applied. For each in situ hybridization image the score ‘highest expression’ was assigned to the segments giving the strongest signal. The score ‘no expression’ was assigned to segments without any detectable signal and the score ‘intermediate expression’ was assigned to segments giving an intermediate intensity signal.

The in situ hybridization data are stored in a relational database (MySQL) and it can be queried using a single web page divided into three frames—a search frame, a result frame and the crypt/villus (CV) navigator frame. The search frame allows text-based queries using gene names or descriptive text. Pressing the ‘search’ button with an empty search box generates a list of all the genes in the database. In addition, pressing either the ‘Crypt specific genes’ or the ‘Villus specific genes’ buttons in the search frame generates lists of crypt or villus specific genes, i.e. genes that are only expressed in the crypt (segments 1 and 2) or villus (segments 3 and 4), respectively. The result frame can display three kinds of results: gene list display, single gene display and help display. Gene list display provides a list of genes fulfilling the search criteria entered in the search frame. From the list a link leads to the single gene display in the result frame. The single gene display provides information about the expression levels in the main parts of the intestine (the jejunum, the ileum and the colon). In a few references no specific information is available about the origin of the small intestinal tissue used in the experiment and in these cases the expression level is displayed in the column termed ‘no positional informational available’. The single gene display is tightly linked to the CV navigator frame. The CV navigator provides an interactive graphical display of the crypt/villus axis (or the colonic axis) and it displays the relative expression level for the active column in the single gene display table. Red color represents the highest expression level, green color represents an intermediate expression level and black color represents the absence of any signal. The CV navigator can also be used to provide a list of genes fulfilling a specified expression pattern. Clicking the segments in the graphical display with the pointer sets the specific pattern. Each click changes the color. When the desired pattern has been set, the gene list is displayed in the result frame by pressing the button ‘search pattern’ in the CV navigator frame.

The help pages provide information about the use of the web interface and additional information about the methods used for collecting and scoring the in situ hybridization data.

Our approach demonstrates that collection, scoring and normalization of in situ hybridization data from the literature when combined with a web interface provides a useful tool to extract comparative information, which is otherwise very difficult to collect from the literature.

ACKNOWLEDGEMENTS
We wish to thank professor Niels Keiding and Professor Hans Sjöström for valuable discussions. The work was supported by the Danish Medical Research Council, The Novo Nordic foundation and the Alfred Nielsen and Wife’s foundation.

REFERENCES