INCOGEN

HDFR: High Density Filter Reader
Version 3

"High-Density Filter Analysis in a Flash"

INCOGEN, Inc.
104 George Perry
Williamsburg, VA 23185

Phone: 757-221-0550
info@incogen.com

www.incogen.com

1. INTRODUCTION

HDFR is a Java-based program designed to efficiently and accurately analyze images of high-density filters. The program includes a user-friendly graphical interface (Figure 1) to help users quickly complete all of the steps involved in the analysis of high-density membranes. HDFR provides an intuitive interface to tasks such as importing an image, selecting or creating a layout, placing and adjusting the grids, identifying clone addresses, performing quantification and expression level studies, viewing data, and verifying and exporting the results. The analysis is based on the user's choice of a pre-existing duplication pattern or a new, custom duplication pattern that can be easily created and implemented by the user.

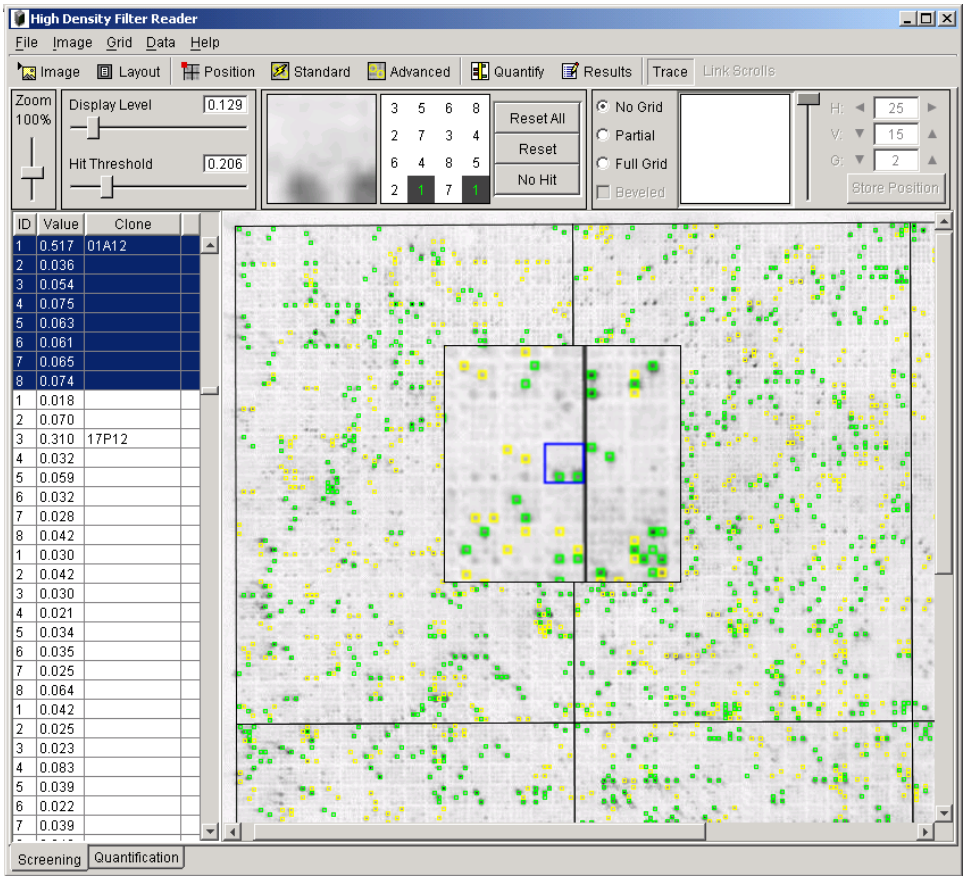


Figure 1: HDFR under Windows 2000 displaying a dynamically magnified portion of the filter. The exact appearance may be slightly different, depending on the operating system and windows manager settings. The location of the buttons, menus, menu items, and other GUI components will be consistent with this screen shot.

2. HIGH-DENSITY FILTER ANALYSIS

The isolation and analysis of DNA clones is an essential part of the molecular analysis of organisms. High-density membranes have been widely used as a resource for screening of clone libraries and as a cost-effective alternative to glass-slide microarrays for expression studies.

The utilization and size of ordered libraries have increased dramatically due to automation. Machines capable of automatically picking colonies, placing them into high-density microtiter plates, and producing high-density ordered filters are commercially available (Genetix, ISS).

High-density, large-insert genomic libraries are available for a number of organisms, such as human, mouse, Arabidopsis, rice, barley, sorghum, and tomato. Filters available from commercial and public sources provide researchers with access to powerful screening technology without the necessary expense associated with the purchase of high-priced specialized equipment.

Clones on high-density membranes are usually spotted in a duplicate spotting pattern for quality control purposes (as well as to facilitate manual reading of the filters). Each location and pattern of a clone on the filter provides a method to identify its address, i.e. the plate and well location. In the past, densities of colonies spotted onto filters have been much lower than the maximum capability of robots because of limitations associated with manual reading of filters. Currently, spotting patterns ranging from 4x4 to 6x6 or higher are commonly used. Using a 4x4 pattern allows the duplicate spotting of 18,432 unique clones per filter (36,864 total colonies), see Figure 2. The 6x6 pattern has 41,472 unique clones per filter (82,944 total colonies). The use of higher patterns is preferred because more clones can be screened in a single probing. However, higher density patterns are not very common because many laboratories still resort to manual filter screening.

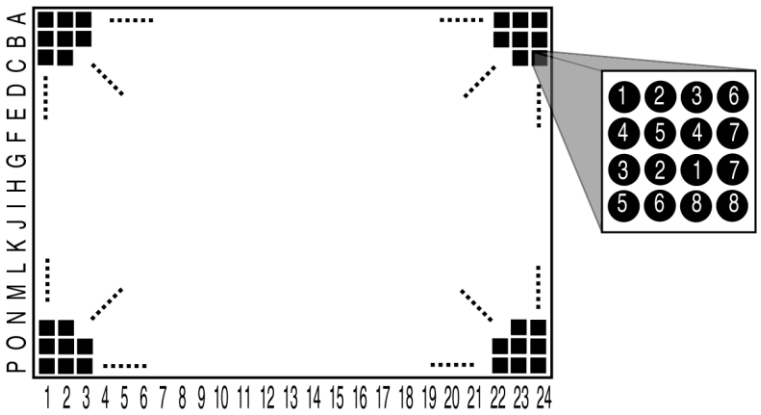


Figure 2: Diagram showing the gridding of colonies in one field of a high-density filter. The colonies are spotted in duplicate in a 4x4 pattern.

The High Density Filter Reader not only greatly increases the efficiency of screening and allows higher-density spotting patterns, but it provides researchers with a means to conduct cost-effective expression studies that would not be feasible without an automated method to analyze the filters.

3. PROGRAM ARCHITECTURE / SYSTEM REQUIREMENTS

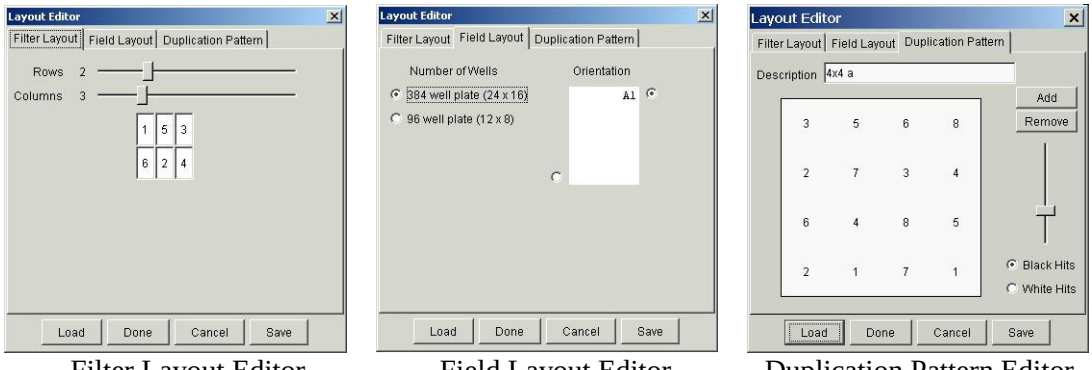
The High Density Filter Reader was written in Java 2 and is available on a variety of computing platforms. Table 1 illustrates the minimum and recommended system requirements for running the software on various platforms.

Operating System	Minimum Requirements	Recommended
Windows 95, 98, 2000, NT, XP	Pentium 200 processor 32 MB of memory SVGA Resolution 10 MB free disk space	Pentium III processor 128 MB of memory 1024x768 resolution
Solaris	Sparc 5 32 MB of memory X-Windows 10 MB free disk space	Ultra 5 128 MB of memory X-Windows
Apple	iMac 300MHz processor 128 MB of memory Mac OS X	G4 processor 128 MB of memory Mac OS X

Table 1: Minimum system requirements and recommended system specifications for running HDFR.

4. DESCRIPTION OF OPERATION

HDFR provides a user-friendly interface to analyze high-density membranes in a high-throughput environment. The program accepts and reads images from a variety of common and proprietary image formats, such as *tiff*, *jpeg*, *gif*, and the Molecular Dynamics *gel* formats. To increase efficiency during the analysis, HDFR stores information relevant to a particular type of filter in a filter layout file. The information stored in a layout file consists of the number of rows and columns of fields, the order of the fields, the plate-type (96 or 384 wells) and orientation, image formats, and duplication patterns. Users can use pre-defined layout files or use the graphical Layout Editor (Figure 3) to quickly and easily define their own custom layout files.



Filter Layout Editor | Field Layout Editor | Duplication Pattern Editor

Figure 3: Example screenshots of the graphical Layout Editor.

Once users load an image and define their filter layout, they are required to position a grid onto the filter image. Several display options are provided to aid the user in accurately positioning the grid (Figure 4). The grid is positioned by using four control points in the corners of the grid and, if applicable, additional control points between fields to adjust the vertical and horizontal inter-field gaps.

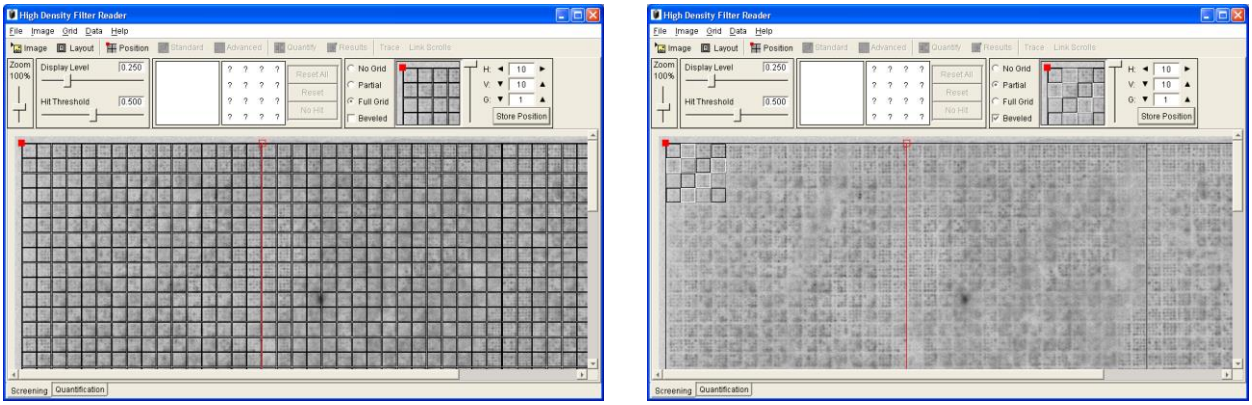


Figure 4: Example display options to aid the user in accurately positioning the grid. The left screenshot shows the "Full Grid" option, while the right screenshot presents the "Partial Grid", "Beveled" option.

The High Density Filter Reader employs an efficient algorithm to perform a standard analysis of the filter image. Using information from the grid and the gap sizes, the filter image is divided into cells of equal size and the intensities of all possible hits are calculated. Background noise levels are eliminated by a weighted average of global and local extreme values. The resulting values correspond to confidence levels to which the user applies a threshold. A standard analysis of an entire filter takes only 2-5 seconds on a Pentium III containing a 500 MHz processor.

If the filter contains many irregularities, users have the option to execute an advanced analysis which is slightly more time-consuming, but will generally perform better than the standard analysis. The advanced analysis compensates for irregularities such as folds in filters or inaccurate placement of the grid.

After the clone addresses have been identified, users can quantify the hybridization events for expression analysis between filters (Figure 5). Multiple filters can be compared to each other at the same time. Results from both types of analyses, clone address identification and quantification, are presented in a graphical, color-coded format superimposed on the filter image, and a tabular, spreadsheet-type format. The combination of both visualization methods allows users to view large sections of data at a glance, as well as quickly jump to detailed analysis results for specific clones of interest. In expression studies, for example, up and down regulation is denoted by green and red outlines, respectively, while the yellow outline signifies the control. Bar graphs provide quick visual representation of relative expression levels, while features such as thresholding allow users to only view expression level ratios greater than a specified value and provide a means to efficiently scan the data. Additional tools, such as multi-level thresholding, interactive zooming and navigational shortcuts allow users to quickly and precisely mine a large volume of data in a short period of time.

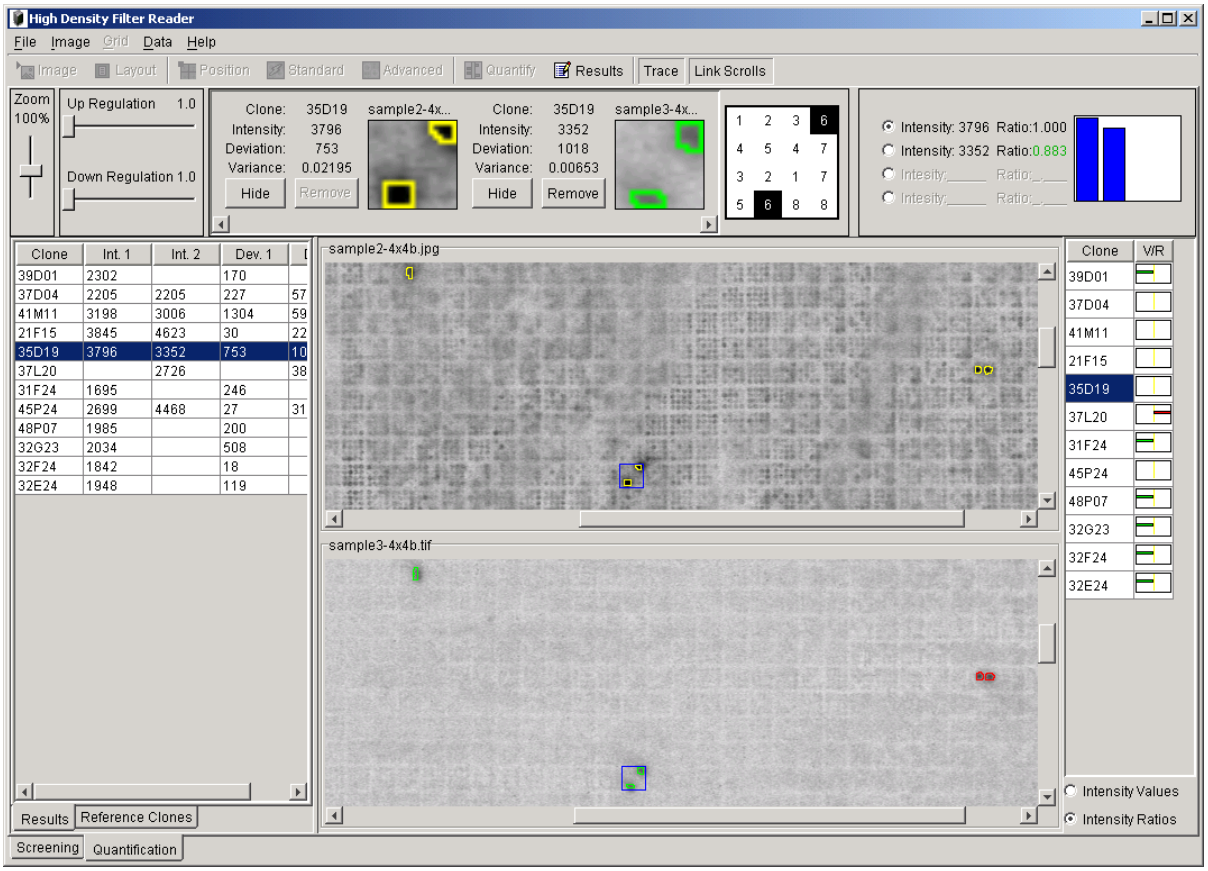


Figure 5: Expression study of three filters. Up and down regulation is denoted by green and red outlines, respectively, while the yellow outline signifies the control. Bar graphs provide quick visual representation of relative expression levels, while features such as thresholding that allow users to only view expression level ratios greater than a specified value provide a means to efficiently scan the data.

Results can be exported in various formats. The default format for exporting data obtained from screening libraries provides a straightforward tab-delimited listing of the clone names, probe names, and hit confidence scores. Users are given options to further customize the output format for specific functions. For library re-arraying purposes, the data can also be exported in a format that can be directly imported into robots such as the Genetix Q Bot. Quantification data is exported with additional values, such as hit intensity and standard deviation.

HDFR is also capable of interacting with the GenePort™ data and discovery management system from INCOGEN, utilizing the underlying database and project tracking mechanisms, as well as linking expression data to sequence information and permitting web access to analysis results.

5. SUMMARY

HDFR provides users with a tool that can dramatically increase the efficiency of many crucial mapping projects and may allow laboratories to carry out high-throughput screening experiments previously thought not feasible because of the inefficiency of manual reading. In addition, by utilizing the quantification functionality of HDFR for high-density membranes, researchers are presented with a very cost-effective alternative to glass-slide microarray technology for performing expression studies.