

CmpSpots - 2D Database Spot Pairing pre-alpha version

Welcome To CmpSpots

<http://www.lecb.ncifcrf.gov/Open2Dprot/CmpSpots/>

and

(<http://open2dprot.sourceforge.net/CmpSpots> mirror)

Note: This pre-alpha level software is under construction and will change.

Not all functionality described is fully implemented or fully debugged.

This documentation will undergo revision as the program develops.

Introduction

CmpSpots is a [Open2Dprot project](#) Java 2D spot data pairing program for finding paired spots between two samples. A [composite quantified protein expression database](#) consisting of protein expression per (spot) for N samples may be constructed from N sample spot lists. CmpSpots is a program to aid in the construction of this type of database by pairing corresponding spots between spot lists based on their 2D (x,y) locations relative to one another.

The spot list data consists of quantified spot with location (x,y) location that varies between samples. Such data could have been created using image segmentation (e.g., 2D electrophoretic gels), derived from non-image data using clustering methods (e.g., 2D LC-MS) to define peptide "spots", or other other methods. For example, if the samples are 2D electrophoretic gels, when we refer to 'sample' we imply gel samples and the spot list of the gel could have been produced by the Open2Dprot [Seg2Dgel](#) spot segmenter and quantification program. Open2Dprot is being designed to handle data from a variety of data sources.

It is a step [4] module in [pipeline analysis](#) for the Open2Dprot project. Spot pairing is performed on data from a reference sample and another sample to be paired with that reference sample. The method uses a set of landmarks defined between the reference sample and the other sample. Landmarks could be created in a [Future] step [3] module in [pipeline analysis](#) operation.

The program may be run either interactively (-gui) with a graphical user interface (GUI) or under an OS shell command to implement batch (-nogui). In the former, after the segmentation is finished, the user has the option of interactively viewing the spot pairing. They may adjust parameters and re-pair the data.

You can currently download the pre-alpha version and install it on your computer. Currently, CmpSpots is hardwired to start with the demo samples and with the -gui switch. The remainder of this home page contains links to some screen shots of the interactive GUI. The Web site contains some initial (rudimentary) documentation.

See the [Reference Manual](#) for details. You read about [downloading and installing](#) the program on your computer. The source code will be put onto [open2dprot.sourceforge.net](#) when it is a bit more stable - currently it undergoing major refactoring.

Examples - samples of screen shots

To give the flavor of running the spot pairing, we provide a few screen shots of the graphical user interfaces and some results. You can these images in the list below or [view all of the screen shots](#) in a single Web page.

- [Initial user interface Report Window](#)
- [Initial command line options tool](#)

Please [contact us](#) with suggestions and comments. If you make interesting changes in the source code, please send us a copy and describe your changes so we can merge them in the released version.

CmpSpots is a contributed program available at [Open2Dprot.
sourceforge.net/CmpSpots](http://Open2Dprot.sourceforge.net/CmpSpots)

Revised: 11/07/2005

[Contact
us](#)

Powered by

SourceForge
Logo

Table of Contents

CmpSpots

[Home](#)

[Introduction](#)

[Overview \(PDF\) \(2\) \(6\)](#)

[Reference Manual\(PDF\)](#)

[Contributors](#)

[Reference manual](#)

[License](#)

[PDFs](#)

[Tutorials](#)

[Demos & screen shots](#)

[Version](#)

[Downloads](#)

[Status](#)

[Revision history](#)

[FAQ](#)

[References](#)

Source code

[Project summary](#)

[CVS access](#)

[Browse CVS](#)

[CmpSpots Javadocs](#)

[O2Plib Javadocs](#)

[Files mirror](#)

Other Web resources

Hosted at Open2Dprot

open2dprot.sourceforge.net

SourceForge

Logo

Powered by

[Contact us](#)

Revised: 05/19/2006

CmpSpots Program Contributors

The current CmpSpots program is the result of a collaborative effort with contributions from various groups. The current Java program is partly based based on some of the code refactored from the C/Unix GELLAB-II 2D gel spot pairing program (cmpgl2, 1993 version). That in turn had been refactored from a DEC-10/SAIL language version. The new Java version also includes code from the open source [Seg2Dgel](#), [Flicker](#), and [MAExplorer](#) programs.

Many individuals have been involved over the years in its development both as direct collaborators running 2D gel experiments as well as in offering suggestions for improvement including: Lew Lipkin, Eric Lester, Jim Myrick, Mary Robinson, Carl Merril, Heinz Busse, Peter Sonderegger, Pete Rogan, Ye Chung Wu, Steve Aley, Mike Alley, Rob Ashmore, Kyle Upton, Jeff Orr, Mort Schultz, George Carman, Mark Miller, Arthur Olson, Norman and Leigh Anderson, Sam Hanash, Jian Li, and many others in the electrophoresis community. Published descriptions of the algorithms for the [GELLAB-II](#) program on the Open2Dprot Web site and are in the [References](#) section of this Web site.

The core group developing the CmpSpots program source currently consists of

- * Peter Lemkin
- * Greg Thornwall
- * Jai Evans

[Contact us](#)

CmpSpots is a contributed program available at [open2dprot.sourceforge.net/CmpSpots](#)

Revised: 04/20/2005

Table of Contents

CmpSpots

[Home](#)

[Introduction](#)

[Overview \(PDF\) \(2\) \(6\)](#)

[Reference Manual\(PDF\)](#)

[Contributoyrs](#)

[Reference manual](#)

[1. Introduction](#)

[Project folder](#)

[Input spot list files](#)

[Paired spot-list file](#)

[Paired-spot viewer](#)

[2. Algorithm](#)

[Primary pairing](#)

[Secondary pairing](#)

[3. Running CmpSpots](#)

[Computing windows](#)

[Local file folders](#)

[Command line args](#)

[Options window](#)

[Updating CmpSpots](#)

[Increasing memory](#)

[4. Command window](#)

[5. Pull-down menus](#)

[Menu notation](#)

[File menu](#)

[Edit menu](#)

[View menu](#)

[Pair menu](#)

[Help menu](#)

[6. Install CmpSpots](#)

[Running Seg2Dgel](#)

[Requirements](#)

[Files downloaded](#)

[7. List of switches](#)

[Switch examples](#)

[List of -debug bits](#)

[8. Demonstrations](#)

[Screen shots list](#)

[All screen shots](#)

[Report output](#)

[Output file - XML](#)

[Output file - SPF](#)

[Batch processing](#)

[9. References](#)

[License](#)

[PDFs](#)

[Tutorials](#)

[Demos](#)

[Version](#)

[Downloads](#)

[Status](#)

[Revision history](#)

[FAQ](#)

[References](#)

Source code

[Project summary](#)

[CVS access](#)

[Browse CVS](#)

[CmpSpots Javadocs](#)

[O2Plib Javadocs](#)

[Files mirror](#)

Other Web resources

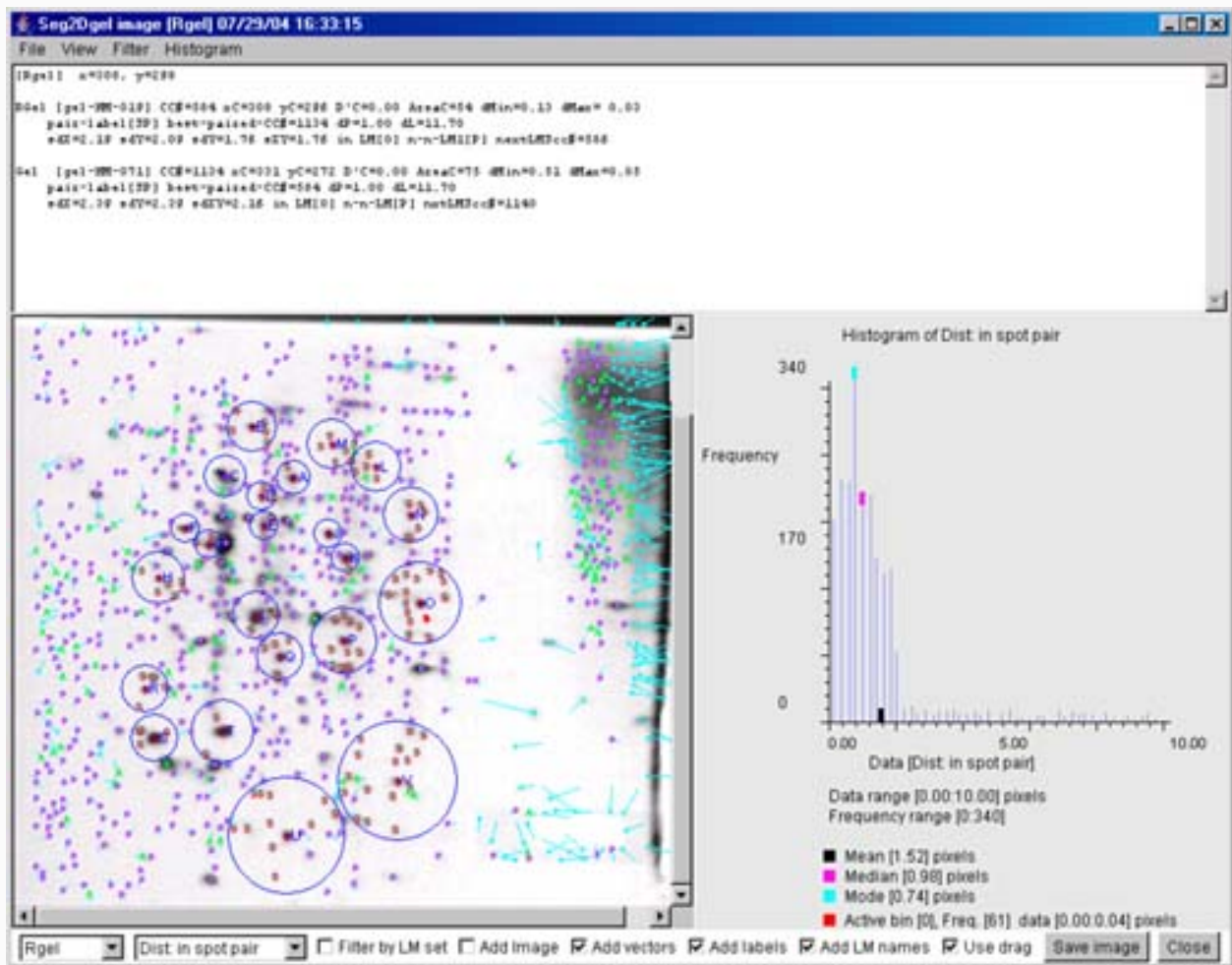
Hosted at Open2Dprot

open2dprot.sourceforge.net

Powered by 

[Contact us](#)

Revised: 05/19/2006



CmpSpots Reference Manual

The following description will tell you what the options are for the CmpSpots and how to run the program - both interactively and from the command line. If you interact with it through the windowing system, then Figure 1 shows a screen view of the CmpSpots reporting window.

CmpSpots Report Window

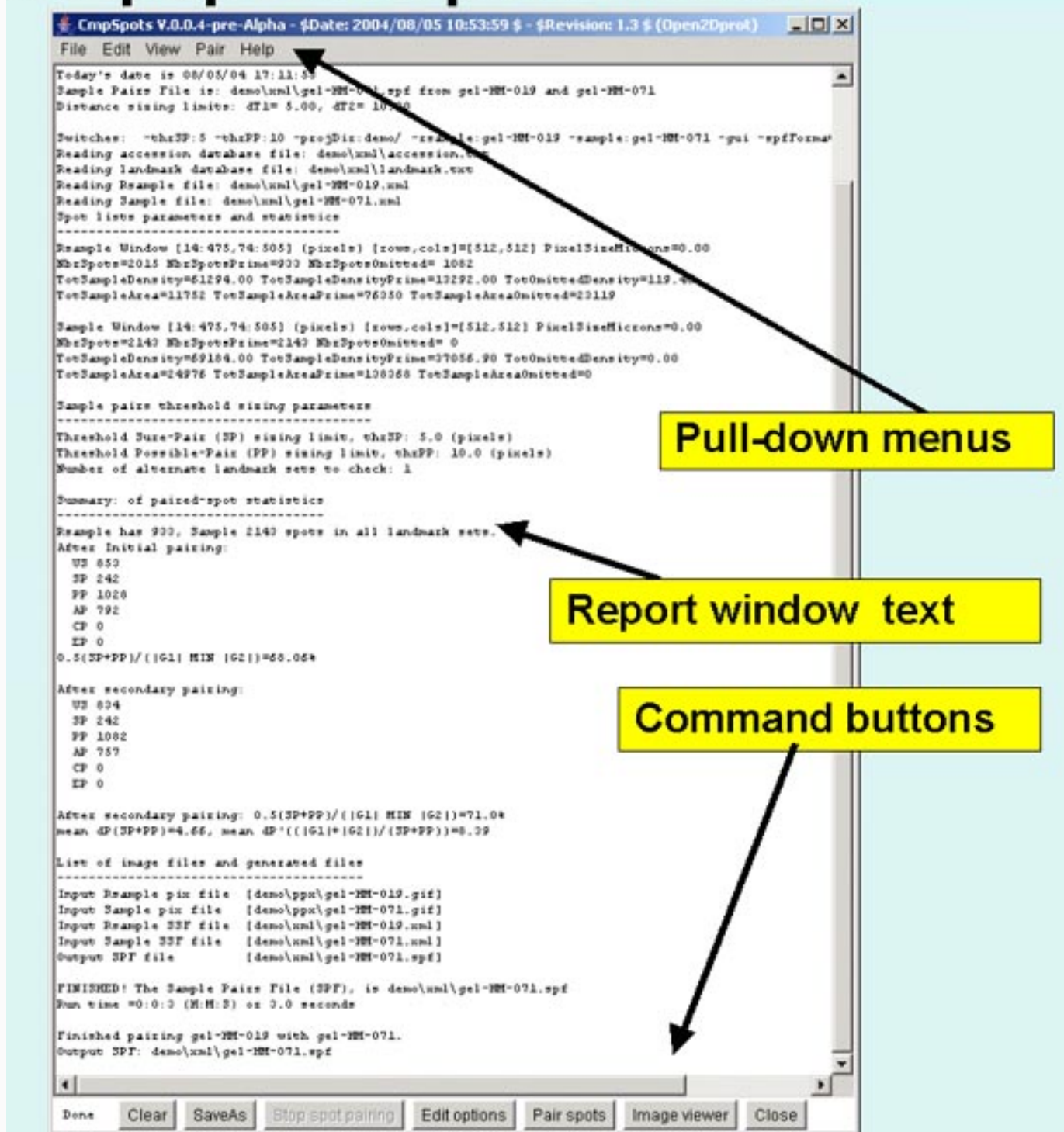


Figure 1. Screen view of the CmpSpots program Report Window interface This screen shot shows the status of CmpSpots in the Report Window graphical user interface after a spot pairing of the demo samples which are 2D PAGE gels. The reference sample or Rsample is gel-HM-019. The other sample denoted Sample is gel-HM-071) has completed. The pairing threshold parameters are shown at the top with the list of **Switches**. The summary statistics on the spot list pairing is summarized in the window including the run time [Run on a 1Ghz Pentium-4 Windows 2000 system]. The Report window has five [pull-down menus](#) at the top of the window used to invoke menus: [File](#), [Edit](#), [View](#), [Pair](#), and [Help](#) operations. Although, the button commandss at the bottom of the window are also available in the menus, they are replicated as buttons for convenience. The **Clear** button will clear the window. **SaveAs** will let you save the Report Window text in a text file. The **Edit options** button pops up a window to let you edit the options including the sample

name, computing window, sizing thresholds etc. The **Pair spots** button starts the spot pairing on the current options. You can stop the spot pairing by pressing the **Stop spot pairing** button. After you have paired the spot lists, you can review the paired spot data by pressing the **Image Viewer** to popup the image viewing window. A status area appears in the lower left corner and reports the current state of the spot pairing during an analysis. It shows "Done" since the spot pairing had completed.

1. Introduction

CmpSpots is an open source Java 2D spot list matching program for finding and pairing spots between the two quantified spot lists. It is part of the Open2Dprot project (<http://open2dprot.sourceforge.net/>). CmpSpots is derived the the GELLAB-II 2D gel spot pairing program **cmpgl2** described in [3]. CmpSpots has been generalized to other types of samples besides 2D gels. While the original program was written in C, CmpSpots is written in Java, uses XML input and output files and has an optional graphical user interface. This initial open-source CmpSpots program code could be used as the basis for more advance spot pairing methods.

The program may be run either interactively (-gui) or under an OS shell command line interface to implement batch (-nogui). If the default -gui mode is used, after the spot pairing is finished, the user has the option of interactively examining the paired spot data overlayed on the original sample images. The user may also modify the input switch options and save the new options in a "CmpSpots.properties" file in the current project directory when they exit so that the last used options may be used as the default switch options in subsequent running of CmpSpots.

In the analysis of 2D image samples (whether real or virtual), the matching of spots between these samples by pairing spots between their spot lists is difficult to perform manually when there are more than a few spots. Automation is necessary when the samples contain a large number of spots, spots which are at best only locally congruent from sample to sample, when the spots can not be counted on to maintain their "shape" or density and contain little infrastructure on which to build a characterization [3]. There is at best a local congruence between two samples related by some a priori undetermined affine transformation. Some of this description is derived from [3].

Partitioned search - dividing the 2D space into local regions

Aside from the number of spots to be examined, a major problem complicating spot localization may be the global distortions in the samples. However, if we are locally well-behaved data, neighboring spots in one sample will likely be neighbors in another gel but the intervening distances between them will change to to a greater degree. The spot pairing problem can be handled various ways.

One method would be to warp (using a non-linear affine transform) one sample image to the geometry of the reference sample image. Then when these images are analyzed by quantifying segmented spots, the spots will be congruent. This method won't work if there are no images (e.g., the 2D spot lists are derived from non-image data (e.g., 2D LC-MS data).

Another method would do a similar warping, but using only the spot lists themselves.

CmpSpots method: A landmark driven spot pairing algorithm

An alternative method is to partition the spot lists into local regions and pair spots within these regions. Within a region, spots surrounding a spot that is known to be paired can most frequently be reliably paired between samples. We will call these spots in the center of the regions "landmarks". The method uses an abstract (x,y) space rather than actual images. Spot (x,y) coordinates are mapped from the sample to the reference sample. This is done, for each landmark, by assigning the landmark spots in the two samples, the cartesian coordinate (0,0). Then using simple linear vector arithmetic, the set of sample spots coordinates are mapped to the reference sample. A set of landmark spots increases the efficiency of inter-sample spot matching by providing an empirical basis for the partitioning of a virtual sample images into tractable corresponding subregions.

A landmark spot may be defined in various ways. In one current empirical procedure for choosing landmarks, it is a morphologically distinctive spot such that neighboring spots and the landmark spot form a consistent morphological structure. Moreover, this structure should be easily recognized across the set of samples used in an experiment consisting of a number of gels (> 2). The landmark spots are selected to cover the regions of interest of the sample fairly evenly if there is uniform distortion between samples. If there are some regions with a lot of distortion, more landmarks should be used in these regions. Depending of the similarity and distortions between the samples, a few to a larger number may be required. This set of landmarks is called the landmark set.

Landmark region

The landmark region is a 2D space surrounding a landmark spot. It is defined as a fuzzy region having more certainty closer in toward the landmark spot. The effective-radius of certainty R_i for landmark i is a distance defined to be half the distance from landmark i to the nearest neighbor landmark. Figure 2 illustrates the effective-radius concept. Images of a spot within the effective-radius of a landmark set would have a higher probability of being aligned (since the landmarks have "perfect" alignment) than if the spot were outside of this radius. Thus, using this heuristic concept of partitioning the spots by landmark region, it is possible to pair the spots automatically once the landmarks are established. The landmark spots are then compared with the two spot lists from the two Sample Spot-list Files (SSF) and the best segmented spot is used rather than the coordinates manually produced. If no spot can be found for a landmark within specified error bounds (currently the maximum latch distance defined using the `-latchLandmarkSpots: maxLatchDist` distance), the landmark coordinate is used in the pairing process. It is possible to check how reliable the particular spot pairing actually was by backchecking paired spots using the Image Viewer.

The probability of finding the same spot in two samples relative to the aligned images of a closeby landmark is greater than if the entire sample spot space were to be searched. This partitioned search has the added advantage that landmark regions contain an order of magnitude fewer spots than the total sample space. Thus the combinatorics of performing the spot matching is greatly decreased as well.

Landmarks could be supplied either using a manual "landmark definition" interactive program (not described here) or might be generated using methods for discovering a subset of robust paired-spots that might be putative landmarks.

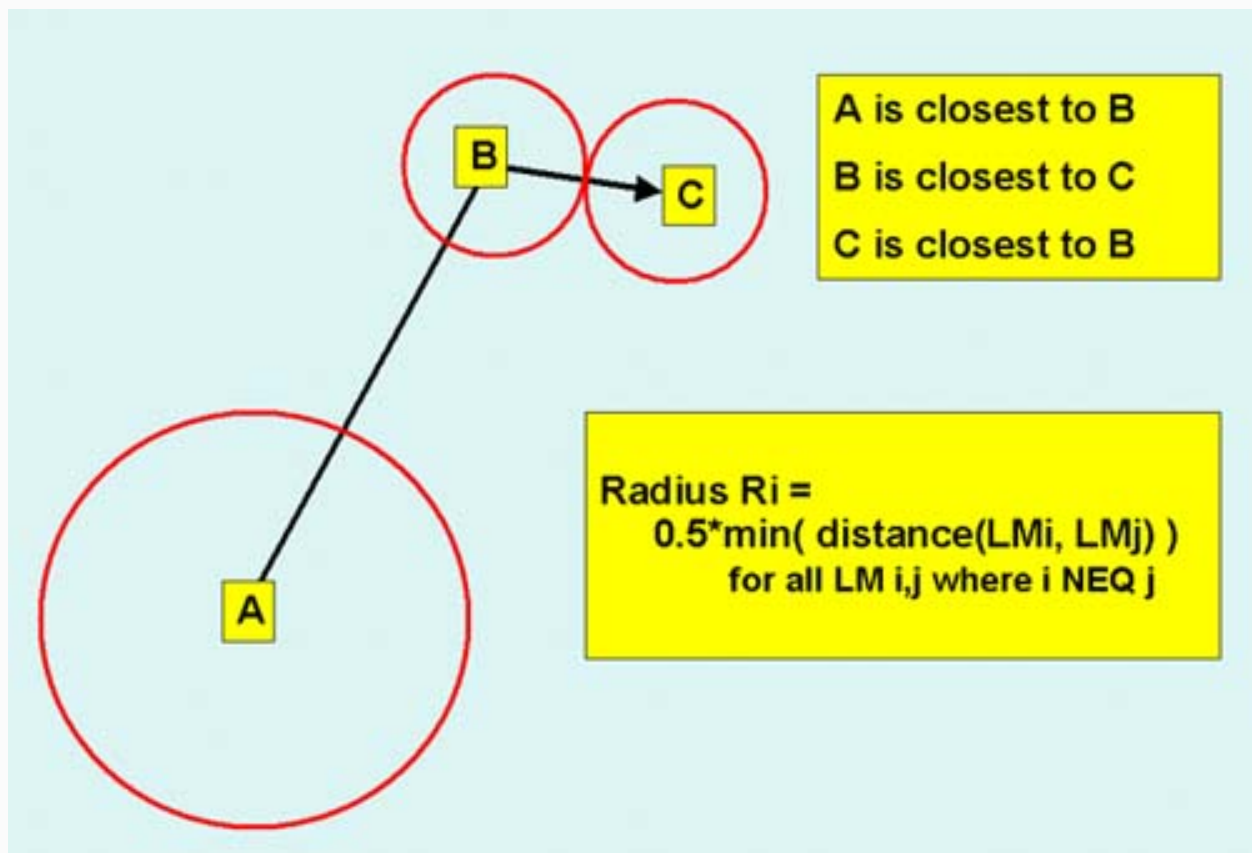


Figure 2. Definition of the effective-radii of a set of landmarks. The effective-radii of certainty of a landmark R_i is $1/2$ the minimum distance from landmark i to its nearest adjacent landmark j . In this example, radius $R_a > R_b$ and $R_b = R_c$. The nearest neighbor of landmark A is B and its next nearest neighbor landmark is C. Estimates of spot pairs within landmark radii are more reliable.

Implementation of landmark-oriented spot pairing between two samples

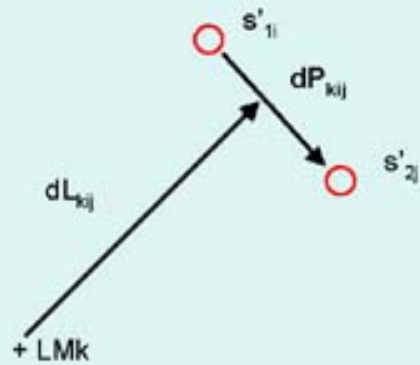
The spot pairing [algorithm](#) is described below. The actual pairing is performed in two passes through the landmark sets data called the [primary](#) and [secondary](#) pairing procedures. Each procedure operates on one landmark set at a time.

In the primary pairing algorithm ([figure 3](#)), the spots are first mapped to the Cartesian coordinate system defined by making the landmark spot (0,0) relative to the origin in the two samples: R_{sample} (G1) and Sample (G2). Each spot in G1 is provisionally paired to the spot that is its nearest neighbor in the projected image of G2. Because of possible asymmetry of the two sets the reverse comparison is also performed so that each spot in G2 is provisionally paired with its nearest neighbor spot in R_{sample} . The nearest neighbor distance is called d_P (pair distance). The distance d_L is the distance from the landmark spot to the mean locus of the two spots in the provisional pair. Two parameter distances are empirically defined: d_{Tsp} and d_{Tpp} . Spots closer than d_{Tsp} are relatively well paired. Spots greater than d_{T2} are very poorly paired and possibly should not be paired. The default values of d_{Tsp} and d_{Tpp} (5 and 10 pixels respectively) were determined empirically, by examination of the nearest neighbor values of several sets of paired samples under a wide variety of conditions. Figure 4 shows various cases which can occur. Four types of pairing labels can be assigned. There are sure pair "SP", possible pair "PP", ambiguous pair "AP" and unresolved spot "US". The primary spot pair labeling assignments are defined in figure 3.

The distance between paired spots is dP_{kij} and is computed after the two spots (s_{1i}, s_{2j}) are mapped to (s'_{1i}, s'_{2j}) in the same sample space relative to landmark k.

$$dP_{kij} = |s'_{1i} - s'_{2j}|.$$

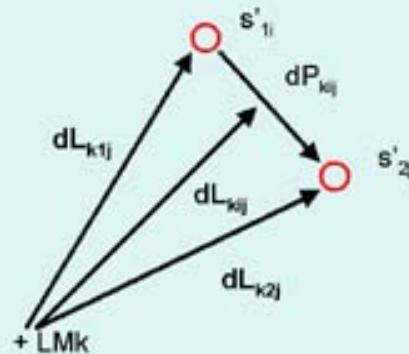
A)



The distance between a spot-pair and the landmark is dL_{kij} and is computed after the two spots (s_{1i}, s_{2j}) are mapped to (s'_{1i}, s'_{2j}) in the same sample space relative to landmark k.

$$\begin{aligned} dL_{k1i} &= \text{distance}(LM_k, s'_{1i}), \\ dL_{k2j} &= \text{distance}(LM_k, s'_{2j}), \\ dL_{kij} &= \min(dL_{k1i}, dL_{k2j}). \end{aligned}$$

B)



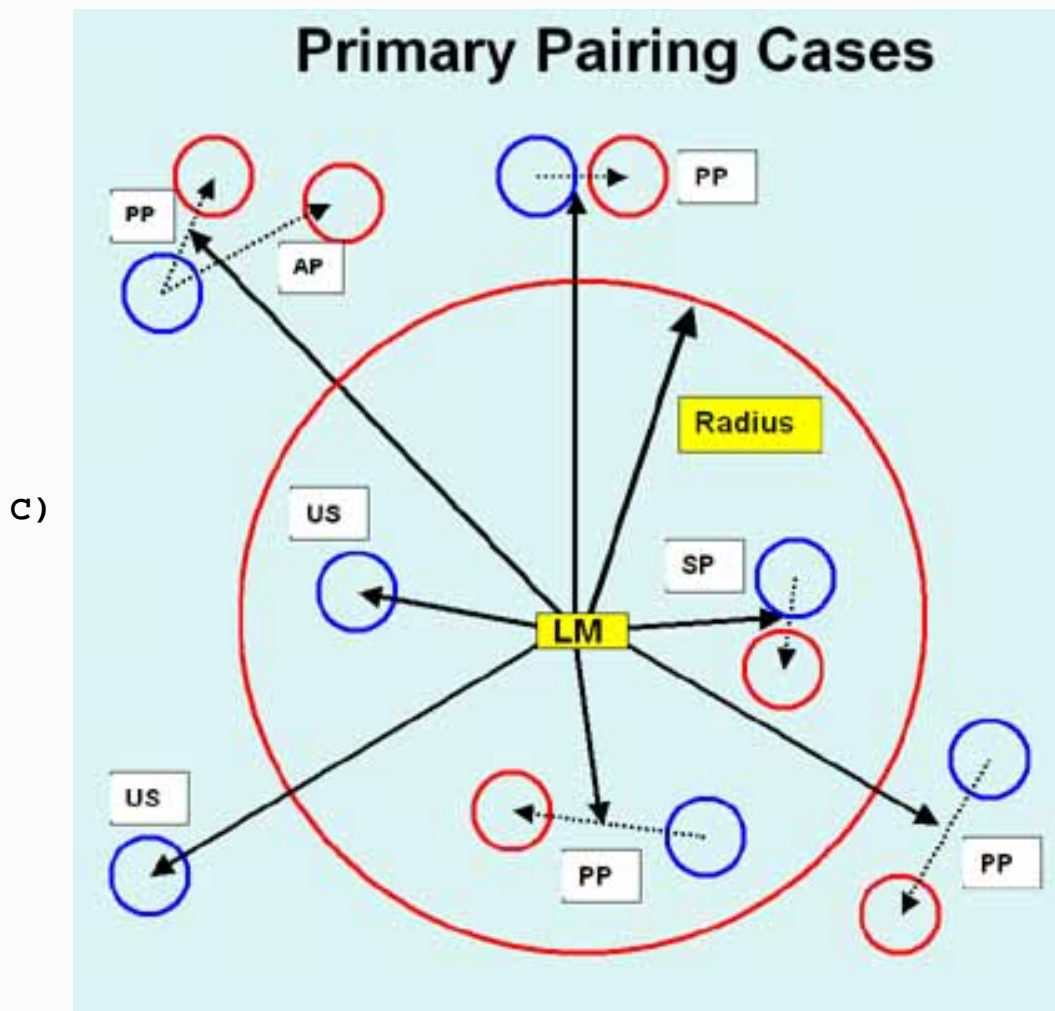


Figure 3. Spot pairing rules. Spot pairing is performed using several features of landmarks and spots in these landmark sets. **A)** Shows the definition of dP or the distance between spots of a spot pair mapped to the same space. **B)** Shows the definition of dL or the distance from a landmark to the spot pair mapped to the same space. **C)** Spot pair primary labeling assignment definitions. Each potential nearest neighbor spot pair in a landmark set has one of four labels: SP - sure pair, PP - possible pair, AP - ambiguous pair, US - unresolved spot. R_k is the effective radius for a landmark set k . The dT_{sp} is the threshold distance for SP spots -**thrSP**: value. The dT_{pp} is the threshold distance for PP spots -**thrPP**: value. [primary pairing](#) is described below. The labeling cases are defined by the following cases:

Case

- [1] US is unresolved spot (no dP) for either sample,
- [2] SP is $dL_2 \leq R_k$ and $dP_2 \leq dT_{sp}$,
- [3] PP is $dL_3 \geq R_k$ and $dP_3 \leq dT_{pp}$,
- [4] PP is $dL_4 \leq R_k$ and $dP_4 \geq dT_{sp}$ and $dP_4 \leq dT_{pp}$,
- [5] PP is $dL_5 \geq R_k$ and $dP_5 \leq dT_{sp}$,
- [6] PP is $dL_6 \geq R_k$ and $dP_6 \leq dT_{pp}$. For the other spot
AP' is $dL_6' \geq R_k$ and $dP_6' \leq dT_{pp}$ and $dP_6' \geq dP_6$.

Secondary spot pairing optimization

A second pass through the data is performed to optimize the spot pairing. Secondary spot pairing can be used to

further resolve AP and US labels in adjacent landmark sets into SP or PP labels which are then placed in either of the two sets. There are four cases: (a) two unresolved spots (US and US), (b) two ambiguous pairs (AP and AP), (c-d) one ambiguous spot (AP) and one unresolved spot (US). The new spot pair is upgraded to a SP or PP pairing label and is put into whichever landmark set has the smallest dL for the putative pair. Each landmark set is tested. Each spot has a list of adjacent landmarks, so these are tested to see if pairing can be optimized for that spot. The [secondary pairing algorithm](#) is described below.

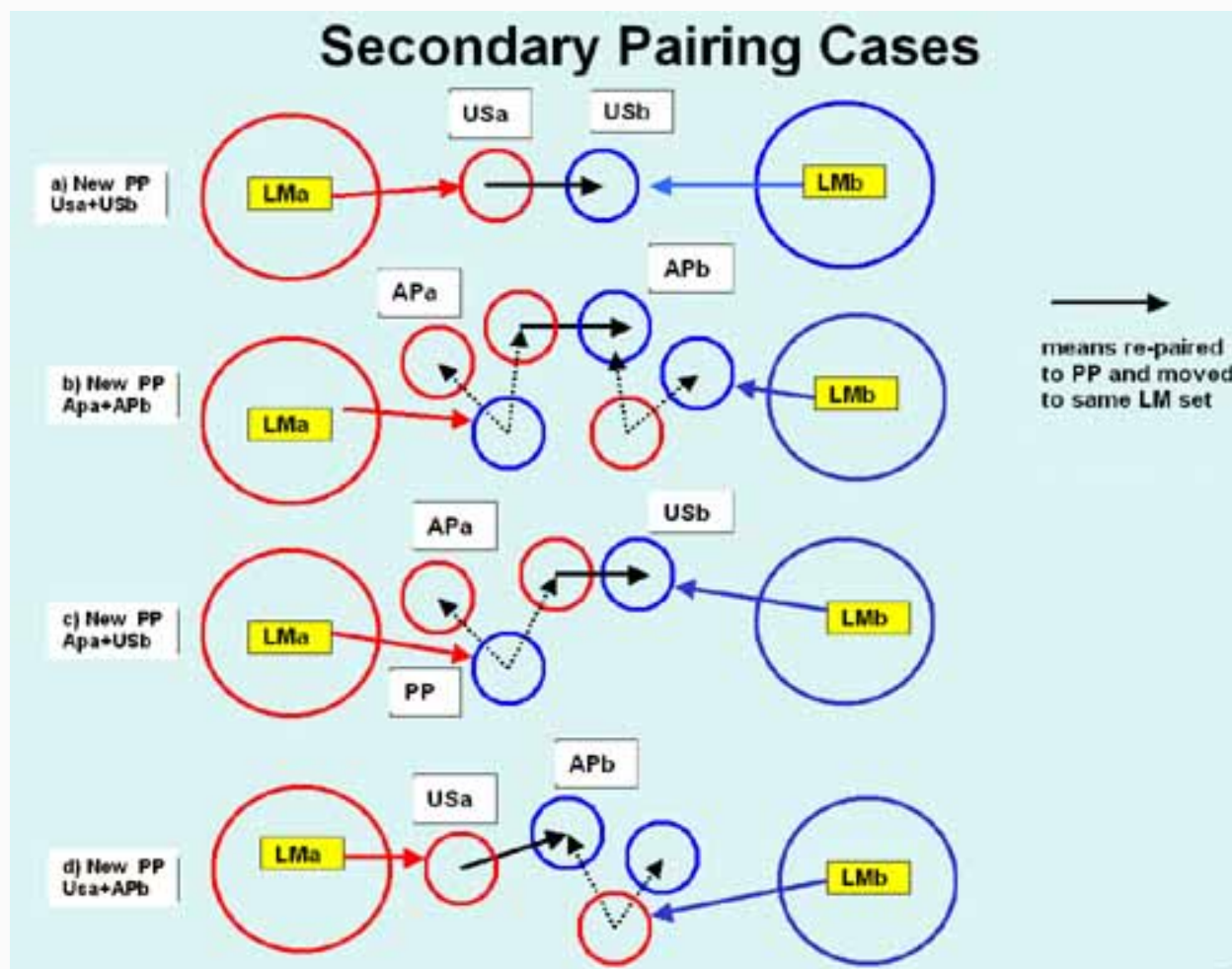


Figure 4. Secondary spot-pairing rules. A second pass through the data is performed to optimize the pairing. Secondary spot pairing can be used to further resolve AP and US labels in adjacent landmark sets into SP or PP labels which are then placed in either of the two sets. There are four cases: (a) two unresolved spots (US and US), (b) two ambiguous pairs (AP and AP), (c-d) one ambiguous spot (AP) and one unresolved spot (US). The new spot pair is upgraded to a SP or PP pairing label and is put into whichever landmark set has the smallest dL for the putative pair. Each landmark set is tested. Each spot has a list of adjacent landmarks, so these are tested to see if pairing can be optimized for that spot.

Project directory structure for Open2Dprot and CmpSpots

All Open2Dprot programs assume a project directory structure. This must exist for the program to proceed. You can either create the structure prior to running any of the programs or you can create it on the fly using the -projDir:user-project-directory. It will lookup and/or create the following sub-directories inside of the user-project-directory.

batch/ directory holding temporary batch files - [NOT USED by CmpSpots]
cache/ directory holding temporary CSD cache files - [NOT USED by CmpSpots]
ppx/ directory holding your original gel input files
rdbms/ directory holding CSD database RDBMS files - [NOT USED by CmpSpots]
tmp/ directory holding generated sample image files
xml/ directory that holds accession DB, landmark DB,
 SSF spot-list files, and generated SPF paired spot list file
 generated by CmpSpots

The use of these directories is discussed in the rest of this document.

Input spot list files

The spot lists to be paired are specified by their sample names (e.g., plasma27). The sample and reference samples to be paired is specified by its image file name using the `-sample` and `-rsample` switches with or without the file extension (e.g., `-sample:plasma27.tif` or `-sample:plasma27`). The file extension is determined by looking up the image in the `ppx/` project subdirectory at run time.

The application looks up the sample in the accession database (in `xml/accession.xml` or as specified using the `-accessionFile` switch) and gets additional information about the sample. The Open2Dprot <http://open2dprot.sourceforge.net/Accession> pipeline module is used for entering samples into the accession database.

[Status: The Open2Dprot Accession module program is not released yet. The accession database could be edited manually as either XML (`accession.xml`), or tab-delimited text (`accession.txt`) with Excel.]

The samples correspond to Sample Spot-list Files (SSF) and should be found in the *user-project-directory/xml/* sub-directory. The SSF format is described in the Open2Dprot [Seg2Dgel](#) program. SSF files may be used by CmpSpots as either XML or full tab-delimited data formats.

There may be images associated with the samples that can be used with the Image Viewer. This is the case if the spots were derived from real images (e.g., 2D gels) or virtual images (e.g., from synthesized 2D LC-MS data), etc. These images may be in TIFF (.tif, .tiff), JPEG (.jpg), GIF (.gif), or PPX (.ppx GELLAB-II) format. TIFF images may be 8-bits/pixel through 16-bits/pixel, whereas JPEG, GIF, and PPX are 8-bit images. Gray values in the image files have black as 0. This is mapped to 0 for white and the maximum pixel value for black.

The input sample image files, if any, are kept in the *user-project-directory/ppx/* sub-directory. This database directory structure is consistent with and is used by the other Open2Dprot analysis pipeline programs.

Landmark database

A small (5 to 25 depending on the rubber-sheet distortion between samples) set of corresponding landmark spots is required for the Sample and Rsample. These are defined in a landmark database (`xml/landmark.xml`) that uses the same sample names as used in the accession database. The Open2Dprot <http://open2dprot.sourceforge.net/Landmark> pipeline module is used for defining landmark pairs for (Rsample, Sample) into the

landmark database.

Output paired-spot list file - the Sample Pairs File (SPF)

The CmpSpots output is a quantified paired-spot list in various ASCII formats including XML and tab-delimited as well as the historical GELLAB-II SPF formats.

The data output file is called the Sample Pairs File (SPF) and is saved in the *user-project-directory/xml/* directory. The generated name the same as the base name of the image file but with a different extension depending on the output format. The possible extensions are: .spf (for ASCII format compatible with GELLAB-II), .xml (XML format), and .txt (tab-delimited format). One of these formats are specified by the -spfFormat:{**F** | **G** | **T** | **X**} command line switch.

[STATUS: -spfFormat:**X** is the default. Note that the XML generated will change with changes in MIAPE.]


1.1 Investigating spot pairing using the paired-spot Image Viewer

Spot pairing may be investigated using the **Image Viewer** button which pops up the Image Viewer window. This lets you select spots to review the pairing data including the quantified spot data. You may add overlays for (subsets of) pairing labels, pairing vectors, landmarks names, landmarks radii. The current selected spot's landmark may be used to filter the spots belonging to that landmark set. There is also a pairing feature histogram that may be used for spot filtering as well. There are four Image Viewer pull-down menus: UL>

- [File menu](#)
- [View menu](#)
- [Filter menu](#)
- [Histogram menu](#)

1.1.1 Image viewer - File menu

These commands are used to change the sample image being displayed.

- **Sample Images**  - specify the reference sample.
 - **Rsample** - display the Rsample image for overlays
 - **Sample** - display the Sample image for overlays
 - **LM Rsample** - display the Rsample image for viewing the landmark half radii circles
 - **LM Sample** - display the Sample image for viewing the landmark half radii circles
- ☒ **Log text to Report Window** - copy all text output from the Image View to the Report window.
- **Save image** - save image to a GIF file.
-
- **Quit** - exit the program when in GUI mode.


1.1.2 Image viewer - View menu

These commands are used to change the sample image overlays being displayed.

- ☒ **Add image overlay, else white** - add overlay
 - ☒ **Add pairing vectors overlay** - add overlay
 - ☒ **Add pairing labels overlay** - add overlay
 - ☒ **Add landmarks overlay** - add overlay
 - ☒ **Add landmark radii overlay** - add overlay
-
- ☒ **User mouse drag to select spots** - select spots by dragging the mouse over spots rather than clicking on them.
-
- ☒ **Report all spot pairs for selected spot** - report in the text area all paired spots incling all AP labeled spot pairs for the selected spot.





1.1.3 Image viewer - Filter menu

These commands are used to change the histogram filter.

- ☒ **Filter spots by histogram bin** - to enable histogram filtering
 - ☒ **Test spot data GEQ histogram data** - otherwise use data less than the histogram bin data you selected.
-
- **Filter spots by pairing label**  - select histogram filter to use
 - ☒ **Add spots with Sure-Pair (SP) label** - filter by SP labels
 - ☒ **Add spots with Possible-Pair (PP) label** - filter by PP labels
 - ☒ **Add spots with Unresolved-Spot (US) label** - filter by US labels
 - ☒ **Add spots with Ambiguous-Pair (AP) label** - filter by AP labels
-
- ☒ **Filter spots in current landmark set** - else use all spots in all landmark sets

1.1.4 Image viewer - Histogram menu

These commands are used to change the histogram filtering criteria.

- Histogram data** Submenu - how to analyze the histogram data
-  **Dist between spots in a pair** - filter by dP or the distance between the two spots after mapped to the same landmark
 -  **Dist between pair and LM** - filter by dL or distance between the center of the spot pair and their landmark
 -  **Labels SP,PP,US,AP** - filter by pairing labels labels
 -  **Landmark set size** - filter landmark set size

2. CmpSpots spot pairing algorithm

The spot pairing algorithm is described in two parts. The first part describes the global processing operations.

The second part describes individual spot pairing operations.

Global processing algorithm

1. The program first loads a sample accession file that contains the names of all samples in the database.
2. The program then loads the two sample SSF spot-list files after verifying them with the accession database. One for the Reference sample and one for the sample being paired. We refer to these as "Rsample" and "Sample".
3. It then loads the set of landmarks for the (Rsample,Sample) pair of samples.
4. The landmark effective radii are computed by finding the distance from each landmark to its closest landmark (for both the Rsample and the Sample). Then a landmark's radius is estimated as 1/2 the minimum of these minimum distances for that landmark (for Rsample and Sample). [Spot pairs inside of this radius can become SP pairs, otherwise they can only become at best PP pairs. This is not counting AP pairs which can occur both inside and outside the radius.]
5. The spots for each sample are assigned to their closest landmark and the spots associated with each landmark are called its landmark set. Each spot also tracks a list of next-nearest landmarks used in secondary pairing.
6. For each landmark, spots in its associated landmark set are optimally paired with spots in its corresponding landmark set in the other sample. Mutually-paired spots are assigned Sure-Pair (SP) and Possible-Pair (PP) labels if distances (described below) are within threshold ranges; The case of N:1 or 1:N paired spots between samples are assigned Ambiguous-Pair (AP) labels (i.e., N spots in one sample are optimally paired with 1 spot in the other sample); Unresolved-Spots (US) are what remains.
7. Secondary pairing is attempted for adjacent landmark sets. For each landmark, we attempt to re-pair or pair spots as SP or PP or (AP and US) with their next-best landmark set. If the pairing is better than the current pairing, we move the spot and upgrade the pairing to a SP or PP. The pairing quality labels are ordered as: SP > PP > (AP or US).
8. Finally it writes out the paired-spot list in the format specified by the -spfFormat switch option. If images are to be saved in disk files, these are also written out at this time.

2.1 Spot pairing within a landmark set - algorithm

This describes the primary spot pairing *within* a given landmark set j . This algorithm is applied for each landmark set in turn. See [Figure 3C](#) which illustrates pairing geometry for two spots in the same landmark set between the two samples. The decision rules use the pairing [cases](#) listed in that figure's legend. Note the symmetry of the algorithm is such that each sample views the other sample the same way.

1. For each spot i in the Rsample landmark set j find the spot k in the Sample landmark set j with the

- smallest dP. Then note the dP and Sample spot k in Rsample spot i.
2. For each spot i in the Sample landmark set j set j find the spot k in the Rsample landmark set j with the smallest dP. Then note the dP and Sample spot k in Sample spot i.
3. For each spot i in the Rsample landmark set j and its optimal spot k in Sample landmark set j, apply the case rules and assign the pairing label. Note: for a spot to be a SP or PP, *both* spots must point to each other. If spot k does not point to spot i, then assign a AP label. If dP is > the dTpp threshold, assign the US label to spot i.
4. For each spot i in the Sample landmark set j and its optimal spot k in Rsample landmark set k, apply the case rules and assign the pairing label. Note: since SP or PP spots were paired in the previous step, we don't do it again - instead only look for spots that do not point to each other. If spot k does not point to spot i, then assign a AP label. If dP is > the dTpp threshold, assign the US label to spot i.

Note: the Image viewer used with the -gui option lets the user interactively investigate this data.

2.2 Secondary Spot pairing - algorithm

This describes the secondary spot pairing. Secondary spot pairing is used to further resolve AP and US labels in adjacent landmark sets into PP labels that are then placed in either of the two sets. There are four cases: (a) two unresolved spots (US and US), (b) two ambiguous pairs (AP and AP), (c-d) one ambiguous spot (AP) and one unresolved spot (US). The new spot pair is upgraded to a PP pairing label and is put into whichever landmark set has the smallest dL for the putative pair. The four cases are described in [Figure 4](#). This procedure is applied to each landmark set in the Rsample and then each landmark set in the Sample. Again, the algorithm is symmetric and is applied to each sample for each landmark set.

1. For each spot i in the Rsample landmark set j that has a US or AP pairing label, for all landmark sets m (such that j is not equals to m) do a trial test on each spot k in the Sample that has a US or AP label. If the trial test has a minimum dP \leq dTpp, then the two spots are a candidate for secondary pairing.
2. Then pair the two spots i and k with a PP label and assign the pair to either landmark set j or landmark set m - whichever has a lower dL value. Remove the spot from the landmark set from which it was removed.
3. For each spot i in the Sample landmark set j that has a US or AP pairing label, for all landmark sets m (such that j is not equals to m) do a trial test on each spot k in the Rsample that has a US or AP label. If the trial test has a minimum dP \leq dTpp, then the two spots are a candidate for secondary pairing.
4. Then pair the two spots i and k with a PP label and assign the pair to either landmark set j or landmark set m - whichever has a lower dL value. Remove the spot from the landmark set from which it was removed.

3. Running CmpSpots and specifying parameter options via the command line

The program may be run either interactively (-gui) with a graphical user interface (GUI) or under an OS shell command to implement batch (-nogui) depending on how it was started. In the former case, after the spot pairing is finished, the user has the option of interactively viewing the paired spot data using the Image Viewer. The user may also modify the input switch options and save the new options in a "CmpSpots.properties" file in the current project directory so that it may be used as the default switch options in subsequent running of the program. [Status: the CmpSpots.properties file is not enabled.] All options including the input reference sample

and other sample to be paired are specified via GNU/Unix style switches on the command line (*-switch{optional 'parameters'}*) and its negation as *-noswitch*). However, if GUI mode is used, you can interactively specify the switches and their options. It is assigned previously by software that generated the SSF spot lists.

The computing window region of interest

The computing window is a rectangular region of interest in the SSF spot list (real or virtual) image where data is considered to be valid. Spots in this region should be paired. Any spots outside of this region are ignored.

The computing window is defined as [x1:x2, y1:y2]. You can set the computing window using the **-cw:x1,x2,y1,y2** command line switch. If you have not defined it or it is not defined in the accession database (if the *-accessionFile* option is used), it is defined as [0:pixWidth-1 x 0:pixHeight-1] where the virtual image is of size pixWidth x pixHeight.

[STATUS: The computing window for each sample entry is defined in the accession database. The Accession program will allow users to define the computing window.]

Local Folders and files created and used by CmpSpots

When CmpSpots is first started, it will check for the following folders and files in the installation directory and create them if they can not be found.

- **CmpSpots.properties** - is the default command line switch options to be used (if present). You can edit the default switch options and then save them into the CmpSpots.properties file.

If you specify an image to be semented, it will check whether it is in a ppx/ subdirectory. If not, it will ask you if you want to create a project directory and will then set up the following four directories and copy your image into the ppx/ directory. You can also use the *-projDir:user-project-directory* switch to specify a (possibly new) project directory.

- **batch/** directory holding temporary batch files - [NOT USED by CmpSpots]
- **cache/** directory holding temporary CSD cache files - [NOT USED by CmpSpots]
- **ppx/** directory holding your original gel input files
- **rdbms/** directory holding CSD database RDBMS files - [NOT USED by CmpSpots]
- **tmp/** directory holding generated sample image files
- **xml/** directory that holds accession DB, landmark DB, SSF spot-list files, and generated SPF paired spot list file generated by CmpSpots

CmpSpots command-line arguments switch usage

The command line arguments usage is:

```
CmpSpots -rsample:Reference-sample -sample:sample [< optional switches >]
```

The complete [list of switches](#) is given later in this manual and as well as some [examples of typical sets of](#)

[switches](#). The user defined default switches may be specified as a resource string 'CmpSpots.properties' file saved in the project directory. For example:

```
CmpSpots -rsample:gel-HM-19 -sample:gel-HM-071 -thrSP:5 -thrPP:10 -project:demo/ -gui
```

Options wizard window for setting the command line switches

If you invoke the **Edit options** button in the Report window (or from the Edit menu), it will popup an options wizard shown in Figure 5 to let you set or change the switch options and then to save these as the new default switch options. The default is saved in the CmpSpots.properties file when you exit program.

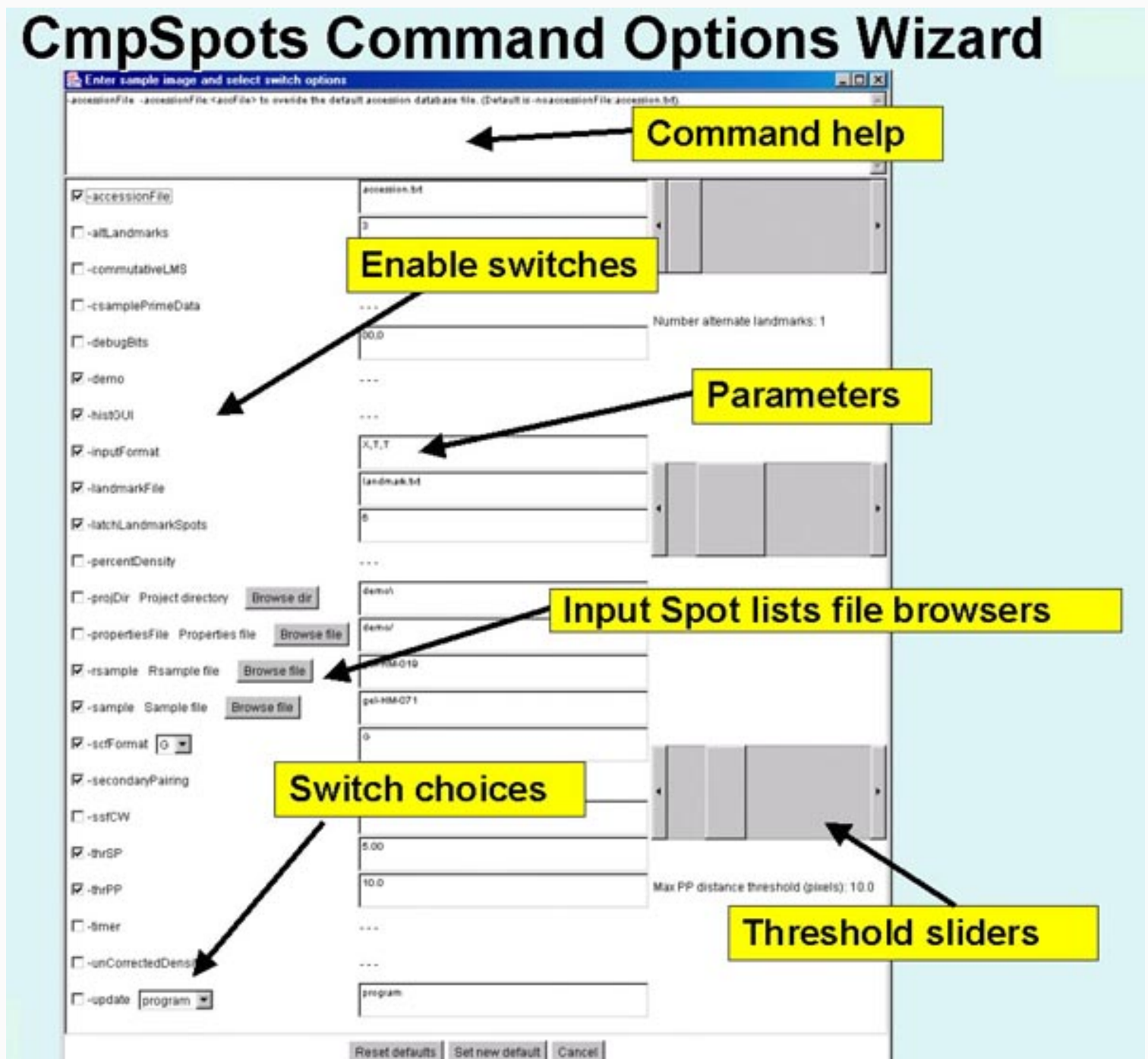


Figure 5. Screen view of the popup options wizard window for setting the command line switches, parameter and specifying input samples to be paired. All of the switches are available in the scrollable window. Switches are checked if they are enabled and if the switch requires a value, the current value is shown in the data entry window to its right. On the right are several threshold sliders for the the upper sizing values for several parameters including **-thrSP:dTsp**, **-thrPP:dTpp**, **-latchLandmarkSpots:maxLatchDist**, and **-altLandmarks:nbrAltLMs**. In the middle, are several Browse buttons to use for specifying a different samples (**-rsample:** and **-sample:**), and directories. Clicking on any switch will show a short help message associated with that switch at the top of the window. Pressing the **Set new default** button will pass the new options values back to CmpSpots. Note: for this to take effect, you must exit and then restart CmpSpots. Then to use them, press the **Pair spots** button in the main Report Window.

Updating CmpSpots from the Open2Dprot Web server using -update switch

As new versions of CmpSpots are developed and put on the Web server, a more efficient way of updating your version is to use the -update commands. There are four options:

<code>-update:program</code>	to update the program jar file
<code>-update:demo</code>	to update the demonstration files
<code>-update:doc</code>	to update the documentation files
<code>-update:all</code>	to update all of the above

After updating the program, it should be exited and restarted for the new program to take effect.

Increasing the allowable memory used by CmpSpots

If you are working with very large images that require a lot of memory, you might want to increase the memory available at startup.

If you are using the Windows CmpSpots.exe file or clicking on the CmpSpots.jar file, you can't change the default startup memory.

However, if you are using the CmpSpots.jar in a script using the java interpreter as in the following example which uses the **-Xmx256M** (specifying using 256Mbytes at startup). Change 256 to a larger size if you want to increase startup memory.

```
java -Xmx256M -jar CmpSpots.jar {additional command line args}
```

4. Command and Report Window - the command center

CmpSpots is designed to be used efficiently in a batch mode with minimal command line output. It is also designed to optionally provide a graphical user interface (GUI) which creates a [Report Window](#) that captures a report of the spot pairing output as well as additional output directed to it by the user. There are a set of [pull-down menus](#) as well as a set of buttons for often used functions.

All logged output is sent to the report window in a scrollable text window that may be saved or used for cut and paste operations. A set of command buttons at the bottom of the window are replicates of commands in the menus, but are easier to access. They include the following functions:






- **Status area** - reports the current status of the spot pairing as it is processing the image.
- **Clear** - clear the report screen
- **SaveAs** - save this report screen in a text file (user is prompted for the file name)
- **Stop program** - stop the spot pairing analysis after finishing analyzing the current landmark set.
- **Edit options** - popup an Options Window to let the user set/clear switches, assign switch parameters (if any), adjust the threshold sizing limits, and specify input samples.
- **Pair spots** - rerun the spot pairing on the (possibly changed) current switch options and input samples
- **Image viewer** - popup the Image Viewer window to investigate details of any images created during spot pairing
- **Close** - save the current SPF file if open, save the `CmpSpots.properties` file, and exit the CmpSpots program.

5. Pull-down menus in the Graphical User Interface (GUI)

The menu bar at the top of the Report Window contains five menus.


1. [File menu](#) - to open the samples to be paired and other file operations.
2. [Edit menu](#) - to change various options including command line switches, computing window, and memory size
3. [View menu](#) - to invoke the image viewer used to inspect the images after spot pairing
4. [Pair menu](#) - to run the spot pairing on your sample spot lists.
5. [Help menu](#) - popup Web browser documentation on CmpSpots

Menu notation

In the following menus, selections that are *sub-menus* are indicated by a . Selections prefaced with a  and indicate  indicate that the command is a checkbox that is enabled and disabled respectively. Selections prefaced with a  and indicate  indicate that the command is a multiple choice "radio button" that is enabled and disabled respectively, and that only one member of the group is allowed to be on at a time. The default values set for an initial database are shown in the menus. Selections that are not currently available will be grayed out in the menus of the running program. The command short-cut notation **C-key** means to hold the Control key and then press the specified *key*.


5.1 File menu

These commands are used to open the samples to be paired and other file operations. The current menus and the menu commands (non-working commands have a '*' prefix) are listed below. You can use either the "Edit options" button to popup the Options Window editor to change the input samples or the **(File menu | Open Rsample)** and **(File menu | Open Sample)** commands.

- **Open Rsample** - specify the reference sample.
- **Open Sample** - specify the other sample to be paired with the Rsample.
- **Set project directory** - set the new default project directory and create the new directory and (batch/, cache/, ppx/, rdbms/, tmp/ and xml/) subdirectories if needed. See the `-projDir:new_project-directory` switch that you can use to do this on startup.
-
- **Clear report** - clear the report text area so that you can start a new report that will be easier to cut and paste or do save to a file
- **Save report to a file** - save the report to a text file
-
- **Update**  - update CmpSpots programs and data from open2dprot.sourceforge.net/CmpSpots server
- **CmpSpots all files** - to get the latest files
- **CmpSpots program** - to get the latest program release (CmpSpots.jar) and other version dependent files from CmpSpots server
- **CmpSpots demo files** - get latest demo project data files
- **CmpSpots documentation** - to get the latest documentation (for stand-alone operation)
-
- **Quit** - exit the program when in GUI mode.

5.2 Edit menu

These commands are used to change various defaults. These are saved when you save the state and when you exit the program.

- **Options**  - change the command line options
- **Open options file** - read a new options file (this overrides the previous options file which could be CmpSpots.properties).
- **Save options file** - write the current options to the current options file (default CmpSpots.properties)
- **SaveAs options file** - write the current options to a new file you specify
- **Edit options** - popup the an Options wizard window to to edit the switch options and thier parameters, threshold parameter sliders, and input image file
- ***Reset default options** - reset options to the initial default

5.3 View menu

This menu contains commands to invoke the image viewer used to inspect the images after spot pairing.

- **Image viewer** - popup the image viewer window to inspect the paired spot data created by the spot

pairing overlayed on the sample images.

- ☒ **Add histogram to Image viewer** - to add a histogram of spot pairing data to the Image viewer (do this before starting the Image viewer).

- **Show landmark DB data** - show the landmark DB data for these samples in the Report Window
- **Show spots matching LMS DB data** - showspots matching the landmark DB data for these samples in the Report Window
- **Show radii of landmark sets** - show the radii of effect for the landmark sets for these samples in the Report Window
- **Show pairing statistics** - show the pairing statistics for these samples in the Report Window
- **Show SPF report** - read the Sample Pairs File generated by the spot pairing and display it in the Report Window
- **Show SPF browser** - read the Sample Pairs File generated by the spot pairing and display it in a popup browser.

5.4 Pair menu


This menu is used to run the spot pairing to perform the analysis. These commands also appear as command buttons at the bottom of the Report Window.

- **Pair spots between samples** - start the spot pairing on the current spots lists with the current option settings
- **Stop spot pairing** - abort the spot pairing after the current landmark set is finished being analyzed.

5.5 Help menu

These commands are used to invoke popup Web browser documentation on CmpSpots. Some of the commands will load local documentation in the the GUI report window.

The documentation is kept on the Internet at <http://open2dprot.sourceforge.net/CmpSpots>. Normally, these help commands should pop up a Web browser that directly points to the CmpSpots Web page. If your browser is not configured correctly, it may not be able to be launched directly from the CmpSpots program. Instead, just go to the Web site with your Web browser and look up the information there.

- **CmpSpots Home** - the open2dprot.sourceforge.net/CmpSpots home page
- **Reference Manual**  - this reference manual
 - **1. Introduction**
 - **2. Algorithm**
 - **3. Running CmpSpots**
 - **4. Command and Report window**
 - **5. Menus**
 - **6. Installing CmpSpots**
 - **7. Command option switches**
 - **8. Demonstrations**
 - **9. References**

- **Sample screen shots** - examples several screen shots
- **Status** - status of the program bugs and future options
- **Revision history** - show latest CmpSpots release and bugs that were fixed
- **FAQ** - frequently asked questions
- **Contributors** - to the spot pairing program sub-project
- **PDF files** - additional documentation including this manual

-
- **Local PDF manual** - popup a PDF browser on the local copy of the PDF document that is also available on the Web site
-

- **Open2Dprot home** - home page for the Open2Dprot project
- **License** - the open source license
- **About** - display information about the program

6. Downloading, installing and running CmpSpots

The installation packages are available for [download](#) from the SourceForge Files mirror. Look for the most recent release named "CmpSpots-dist-V.XX.XX.zip". These releases include the program (both as Windows .exe file and a .jar file), required jar libraries, demo data, Windows batch and Unix shell scripts. Download the zip file and put the contents where you want to install the program. Note that there is a CmpSpots.exe (for Windows program). You might make a short-cut to this to use in more easily starting the program. Alternatively, you can use the sample .bat and .sh scripts to run the program explicitly via the java interpreter. Note that this method assumes that you have Java installed on your computer and that it is at *least* JDK (Java Development Kit) or JRE (Java Runtime Environment) version 1.5.0. If you don't have this, you can download the latest version free from the java.sun.com Website.

6.3 Running CmpSpots

There are several ways to run the program. On Windows, you can start CmpSpots by clicking on the startup icon shown in Figure 6 below. For Unix systems including MacOS-X, you can start CmpSpots from the command line by running the CmpSpots.jar file. If your computer is setup to execute jar files, just type the jar file. In both systems, you can specify additional command line arguments in Windows .bat and unix .sh scripts (see [demo examples below](#)).



Figure 6. Startup icon for CmpSpots.exe in Windows. Clicking on the icon starts CmpSpots. To start CmpSpots, click on the startup icon shown in Figure 6 below - or you can run the [demo-CmpSpots.bat](#) script. For Unix systems including MacOS-X, you can start CmpSpots from the command line by clicking on the CmpSpots.jar file or using

the [demo-CmpSpots.sh](#) script. These two scripts run the program in batch. There are also GUI versions of the two scripts [demo-CmpSpots-GUI.bat](#) and [demo-CmpSpots-GUI.bat](#) that will pop up a user interface. You could make short-cuts (Windows) or symbolic-links in Unix to make it easier to start.

6.4 Requirements: minimum hardware and software requirements

A Windows PC, MacIntosh with MacOS-X, a Linux computer or a Sun Solaris computer having a display resolution of at *least* 1024x768. We find that a 1024x768 is adequate, but a 1280x1024 screen size much better since you can see the Popup Report window, Options window, and Image Viewer window at the same time. At least 30 Mb of memory available for the application is required and more is desirable for comparing large images or performing transforms. If there is not enough memory, it will be unable to load the images, the transforms may crash the program or other problems may occur.

An Internet connection is required to download the program from the Open2Dprot CmpSpots Web site. New versions of the program and associated demo data will become available on this Web site and can be [uploaded](#) to your computer using the various (**File | Update | ...**) menu commands. We currently distribute CmpSpots so that it uses up to 256Mb. See discussion on [increasing memory](#).

6.5 Files included in the download

The following files are packaged in the distribution you install. you can periodically a (File | Update from Web server | ... program) menu command to update the files from the open2dprot.sourceforge.net Web server.

- **CmpSpots.jar** is the Java Archive File for CmpSpots that is executed when you run CmpSpots.
- **CmpSpots.exe** is the Windows executable file that packages CmpSpots.jar with a java startup. This was packaged using the open-source launch4j (launch4j.sourceforge.net) program.
- [Open2Dprot-SSF.dtd](#) is the XML .dtd file used to for the SSF .xml file.
- [Open2Dprot-SPF.dtd](#) is the XML .dtd file used to generate the SPF .xml file if the -spfFormat:X switch option is used. Note: this is only in generating the SPF file if -dtd switch is set.
- [Open2Dprot-Accession.dtd](#) is the XML .dtd file that can be used with the accession database file Accession.xml.
- [Open2Dprot-Landmark.dtd](#) is the XML .dtd file that can be used with the landmark database file Landmark.xml.
- [PDF/fullCmpSpotsDoc.pdf](#) is a copy of the CmpSpots Web site for use as a reference manual when using the program off-line. It may not be as up to date as the one on the web site but you can do a (File | Update from Web server | CmpSpots documentation) menu command to update it.
- **jai_core.jar** is the core Java runtime from SUN's Java Advanced Imaging (JAI) at [java.sun.com](#)
- **jai_codec.jar** is the JAI tiff file reader from SUN's Java Advanced Imaging JAI at [java.sun.com](#)
- **xercesImpl.jar** is the Xerces 2 XML SAX/DOM XML library from Apache [xml.apache.org](#)
- **xml-apis.jar** is the Xerces 2 XML SAX/DOM XML library from Apache [xml.apache.org](#)
- is contains the Mozilla.1.1.html and LEGAL.txt files for CmpSpots and the release-license-jai_files for the Java Advanced Imaging .jar files used by the TIFF file reader.
- **demo/** is a demonstration project directory containing the the following subdirectories:
 1. **ppx/** directory holds sample image files (4 images)

2. **tmp/** directory holds generated image files

3. **xml/** directory that the accession DB, landmark DB, input SSF and SPF output files

You can do a (File | Update from Web server | CmpSpots demo files) menu command to update it.

7. List of the command line switches

The command line usage is:

```
CmpSpots -rsample:reference-sample -sample:sample [< optional switches >]
```

where the order of arguments is not relevant. In the following list, items in **bold** are specific values which must be used (e.g., for -spfFormat:{**X** | **F** | **T** | **G**}, whereas variable values in *italics* indicate that a numeric value for that variable should be used (e.g., for -thrSP:*t1threshold* it might be -thrSP:5). Some switches have several alternate fixed choices in which case this indicated as a list of bolded items inside of a set of '{...}' with '|' separating the items. You must pick one of the items and do not include the '{}' brackets. Also, do NOT include any extra spaces in the arguments of the switch - it will be counted as if it were another switch.

Command line switches

```
-accessionFile:accFile to override the default accession
    database file. (Default is -noaccessionFile:accession.xml).

-altLandmarks:NbrOfLandmarkSets specifies the number of alternate
    landmarks to check. (Default is -altLandmarks:3).

-commutativeLMS lets you search the landmark database for sample pairs
    where the Rsample and Sample are swapped. (Default is
    -commutativeLMS).

-csamplePrimeData indicates that the SSF data are canonical Csample'
    data created from replicate samples in the Composite Sample
    Database (CSD). (Default is -nocsamplePrimeData).

-debugBits:bits,optLandmarkNbr dumps various conditional
    debugging parameters onto the report window as well as the
    output SPF file. The debugging is active for all landmarks
    unless optLandmarkNbr is specified. The 'bits' are the debug
    bits specified as either octal or decimal and enable particular
    debugging output if the program was compiled with debugging
    enabled. (Default is -nodebug).

-default sets the default switches to a specific configuration:
    -nodemo
    -thrSP:5 -thrPP:10 -inputFormat:X,X,X
    This disables -demo if it was set. (Default -nodefault).
```

- demo sets the default switches and sample input sample to a specific configuration. This may be overridden by turning off the -demo switch in the Options Wizard.
- dtd adds the XML DTD file (Open2Dprot-SPF.dtd) in the output XML if -spfFormat:X is set. (Default is -nodtd).
- gui to start the spot pairing with a popup Graphical User Interface rather than in batch mode. This captures messages from CmpSpots. You can then cut and paste the results or save it to a text file. The GUI is also used to change the switch options, re-run the spot pairing and view images after each analysis. (Default is -nogui).
- histGUI will bring the the dynamic histogram graphical user interface when the Image Viewer is requested. (Default is -histGUI).
- inputFormat:ssfMode,accMode,lmsMode defines the input formats for the Sample Spot-list File (ssfMode), accession database file (accMode), and landmark database file (lmsMode). Where mode is: **F** for full tab-delimited data that includes the 3 files: list of spots, parameters, and statistics. The **T** for tab-delimited (.txt) data, and **X** for XML (.xml) data. (Default is -noinputFormat:X,T,T).
- landmarkFile:lmsFile to override the default landmark database file. (Default is -nolandmarkFile:landmark.xml).
- latchLandmarkSpots:maxLatchDist is the maximum distance to use when trying to latch a landmark to its closest spot. (Default is -nolatchLandmarkSpots:6).
- percentDensity to output percent density instead of OD values. (Default is -nopercentDensity).
- projDir:alternate project directory path to specify the project directory to use instead of the default 'demo/' file in the installation directory. (Default is -noprojDir).
- propertiesFile:alternate 'CmpSpots.properties' file to specify the alternate startup properties file to use instead of the default 'CmpSpots.properties' file. (Default is -noproPERTIESFile).
- rsample: is the reference sample.
- sample: is the Sample to be paired with the Rsample.
- spfFormat:{**F** | **G** | **T** | **X**} defines the output format for the

Sample Pairs File (SPF) where mode is: **F** for full tab-delimited data that includes the list of spots but also the same information as the **G** format), **G** for GELLAB-II (.spf), **T** for tab-delimited (.txt) spots-only, and **X** for XML (.xml). (Default is -spfFormat:X). The Open2Dprot-SPF.dtd file is included only if -dtd switch is set.

- secondaryPairing to perform secondary pairing. (Default is -secondaryPairing).
- ssfCW:x1,x2,y1,y2 to ignore spots outside of this window for both SSF input files. This is a debugging tool. (Default is -nssfCW).
- thrSP:dTsp specifies the minimum distance threshold between paired spots for them to be considered a Sure-Pair. (Default is -thrSP:5).
- thrPP:dTpp specifies the minimum distance threshold between paired spots for them to be considered a Possible-Pair. (Default is -thrPP:10).
- timer enables a timer to capture processing times for each step. (Default is -notimer).
- unCorrectedDensity use uncorrected density in the pairing output. (Default is -nounCorrectedDensity).
- update:{**all** | **program** | **demo** | **doc**} specifies that all of the CmpSpots files, the program jar files, the documentation files or the demonstration files should be updated from the Open2Dprot Web server. The program should be exited and restarted after updating the program for this to take effect. (Default is -noupdate).

7.1 Examples of some typical sets of switches

The following shows a few examples of useful combinations of command line switches.

Any case-independent switch may be negated by preceeding it with a 'no' eg. '-notimer'.

The command line syntax used to invoke it is:

```
CmpSpots input-sample-image-file [< opt.-switches >]
```

The following examples using switches might be useful:

```
CmpSpots -rsample:gel-HM-019 -sample:gel-HM-071 -nogui -project:demo/
# Pair two samples into a SPF file using "-spfFormat:X" XML format.
# This is what is normally used in batch mode
```

```
CmpSpots -rsample:gel-HM-019 -sample:gel-HM-071 -gui -project:demo/
# Pair two samples into a SPF file, popup up Report window and then
# use may pop up Image viewer to display sample images with paired spot
# overlays including spot labels, vectors and landmark sets
```

```
CmpSpots -rsample:gel-HM-019 -sample:gel-HM-071 -spfFormat:G
        -project:demo/
# Pair two samples into a SPF file using the GELLAB format.
```

```
CmpSpots -rsample:gel-HM-019 -sample:gel-HM-071 -thrSP:10 -thrPP:20 \
        -altNbrLMS:2 -project:demo/
# Pair two samples as above, but use new thresholds thrSP, thrPP and
        the alternate number of landmarks.
```

7.2 Debug option bits for the -debug switch

The following are the orthogonal octal -debug option code bits. This means you can add them together (in octal) and use that computed octal number (it will also accept decimal). The -debug:*bits,landmarkNbr* option is meant only for serious programmers reading or modifying the source code.

Debugging option bits used with the "-dbug:" command line command

The following are the orthogonal -debug option code bits in octal. This means you can add them together (in octal) and use that number.

	Octal code =====	Methods traced =====
bit:	01 =	assignCodes()
bit:	02 =	assignSSFtoLMS()
bit:	04 =	computeRadii()
bit:	020 =	dumpSPF()
bit:	040 =	findLMSinSSFs()
bit:	0100 =	switchAnalysis()
bit:	0200 =	initialSpotPairing() and convolveG1G2spots()
bit:	0400 =	-free-
bit:	01000 =	printAt()
bit:	02000 =	cvtLMsetSpotsToStr()
bit:	04000 =	cvprintSpot()

```

bit:          010000 = -free-
bit:          020000 = readSSFdata( )
bit:          040000 = readLandmarks( )
bit:          0100000= secondarySpotPairing( )
bit:          0200000 = main( )
bit:          0400000 = dumpLMsetSpots( )
bit:          01000000 = dumpSPFpreface( ) and dumpStatisticsToSPFfile
=====

```

8. Demonstrations

*** BEING REWRITTEN ***

8.1 Examples - samples of screen shots

To give the flavor of running the spot pairing program, we provide a few screen shots of the graphical user interfaces and some images generated by the program.

You can these images in the list below or [view all of the screen shots](#) in a single Web page.

- [1. Initial Report Window](#)
- [2. Initial command-line options tool](#)
- [3. Image Viewer tool](#)
- [3.1 Image viewer - no overlay and no filters](#)
- [3.2 Image viewer - landmarks](#)
- [3.3 Image viewer - landmark sets radii](#)
- [3.4 Image viewer - radii labels no filter](#)
- [3.5 Image viewer - labels no filter](#)
- [3.6 Image viewer - labels filter LM set P](#)
- [3.7 Image viewer - radii labels filter LM set P](#)
- [3.8 Image viewer - vector labels no filter](#)
- [3.9 Image viewer - vector labels filter LM set V](#)
- [3.10 Image viewer - vector labels radius filter LM set V](#)
- [3.11 Image viewer - labels filter LM sets A, B, C, D](#)
- [3.12 Image viewer - histogram filter](#)
- [3.13 Image viewer - histogram filter by dP limit](#)
- [3.14 Image viewer - histogram filter dL limit](#)

8.2 Example - output of the Report Window for a spot pairing

The following Report Window output was generate for the images in the above example.

CmpSpots V.0.0.4-pre-Alpha - \$Date\$ - \$Revision\$ (Open2Dprot)

Today's date is 08/04/04 14:37:34

Switches: -thrSP:5 -thrPP:10 -projDir:demo/ -rsample:gel-HM-019
-sample:gel-HM-071 -gui -spfFormat:X,T,T -spfFormat:G

CmpSpots V.0.0.4-pre-Alpha - \$Date\$ - \$Revision\$ (Open2Dprot)

Today's date is 08/04/04 14:37:34

Sample Pairs File is: demo\xml\gel-HM-071.spf from gel-HM-019 and gel-HM-071

Distance sizing limits: dT1= 5.00, dT2= 10.00

Switches: -thrSP:5 -thrPP:10 -projDir:demo/ -rsample:gel-HM-019 -sample:gel-HM-071 -gui -
spfFormat:X,T,T -spfFormat:G

Reading accession database file: demo\xml\accession.txt

Reading landmark database file: demo\xml\landmark.txt

Reading Rsample file: demo\xml\gel-HM-019.xml

Reading Sample file: demo\xml\gel-HM-071.xml

Spot lists parameters and statistics

Rsample Window [14:475,74:505] (pixels) [rows,cols]=[512,512] PixelSizeMicrons=0.00
NbrSpots=2015 NbrSpotsPrime=933 NbrSpotsOmitted= 1082
TotSampleDensity=61294.00 TotSampleDensityPrime=13292.00 TotOmittedDensity=119.40
TotSampleArea=11752 TotSampleAreaPrime=76350 TotSampleAreaOmitted=23119

Sample Window [14:475,74:505] (pixels) [rows,cols]=[512,512] PixelSizeMicrons=0.00
NbrSpots=2143 NbrSpotsPrime=2143 NbrSpotsOmitted= 0
TotSampleDensity=69184.00 TotSampleDensityPrime=37056.90 TotOmittedDensity=0.00
TotSampleArea=24976 TotSampleAreaPrime=138368 TotSampleAreaOmitted=0

Sample pairs threshold sizing parameters

Threshold Sure-Pair (SP) sizing limit, thrSP: 5.0 (pixels)
Threshold Possible-Pair (PP) sizing limit, thrPP: 10.0 (pixels)
Number of alternate landmark sets to check: 1

Summary: of paired-spot statistics

Rsample has 933, Sample 2143 spots in all landmark sets.

After Initial pairing:

US 853
SP 242
PP 1028
AP 792
CP 0
EP 0

$0.5(SP+PP)/(|G1| \text{ MIN } |G2|)=68.06\%$

After secondary pairing:

US 834
SP 242
PP 1082
AP 757
CP 0
EP 0

After secondary pairing: $0.5(SP+PP)/(|G1| \text{ MIN } |G2|)=71.0\%$

mean dP(SP+PP)=4.66, mean dP'((|G1|+|G2|)/(SP+PP))=8.39

List of image files and generated files

Input Rsample pix file [demo\ppx\gel-HM-019.gif]

```

Input Sample pix file    [demo\ppx\gel-HM-071.gif]
Input Rsample SSF file   [demo\xml\gel-HM-019.xml]
Input Sample SSF file    [demo\xml\gel-HM-071.xml]
Output SPF file          [demo\xml\gel-HM-071.spf]

```

FINISHED! The Sample Pairs File (SPF), is demo\xml\gel-HM-071.spf
 Run time =0:0:3 (H:M:S) or 3.1 seconds

Finished pairing gel-HM-019 with gel-HM-071.
 Output SPF: demo\xml\gel-HM-071.spf

8.3 Examples - part of Sample Pairs File XML format

This is part of a Sample Pairs File to illustrate the type of data available as output. This used the default -
spfFormat:X option with the XML DTD file [Open2Dprot-SPF.dtd](#).

[Open2Dprot-SPF.dtd](#) (if **-dtd** switch used), else

```

<?xml version="1.0" ?>
<SpotList_Pairing>
<Pairing_parameters>
  <date>"05/10/05 11:18:32"</date>
  <Open2Dprot_SPF_Version>"1.5"</Open2Dprot_SPF_Version>
  <Project_directory>"demo\"</Project_directory>
  <Sample_Pairs_File>"gel-HM-071-SPF.xml"</Sample_Pairs_File>
  <thrSP_threshold>5.0</thrSP_threshold>
  <thrPP_threshold>10.0</thrPP_threshold>
  <nbrAltLandmarks>1</nbrAltLandmarks>
  <nbrLandmarks>22</nbrLandmarks>

<Rsample>
  <Sample_Type>"Rsample"</Sample_Type>
  <Sample_Name>"gel-HM-019"</Sample_Name>
  <Simple_FileName>"gel-HM-019.gif"</Simple_FileName>
  <Sample_Pix_FileName>"demo\ppx\gel-HM-019.gif"</Sample_Pix_FileName>
  <Sample_SSF_FileName>"demo\xml\gel-HM-019-SSF.xml"</Sample_SSF_FileName>
  <cw1>14</cw1>
  <cw2>475</cw2>
  <cw1>74</cw1>
  <cw2>509</cw2>
  <Pix_Height>512</Pix_Height>
  <Pix_Width>512</Pix_Width>
  <PixelSizeMicrons>0.00</PixelSizeMicrons>
  <NbrSpots>1588</NbrSpots>
  <NbrSpotsPrime>738</NbrSpotsPrime>
  <NbrSpotsOmitted>850</NbrSpotsOmitted>
  <TotSampleDensity>8838.70</TotSampleDensity>
  <TotSampleDensityPrime>5670.40</TotSampleDensityPrime>
  <TotOmittedDensity>91.90</TotOmittedDensity>
  <TotSampleArea>36158</TotSampleArea>
  <TotSampleAreaPrime>23019</TotSampleAreaPrime>
  <TotSampleAreaOmitted>13139</TotSampleAreaOmitted>
  <PctOmittedToDprimeAcceptedSpots>1.62    </PctOmittedToDprimeAcceptedSpots>
  <Nbr_Spots_Failing_Area_Sizing>
    <nbr_below_t1Area_thr>3316      </nbr_below_t1Area_thr>

```

```

    <nbr_above_t2Area_thr>0          </nbr_above_t2Area_thr>
    <nbr_below_t1Density_thr>4214    </nbr_below_t1Density_thr>
    <nbr_above_t2Density_thr>0       </nbr_above_t2Density_thr>
    <nbr_below_t1Range_thr>0         </nbr_below_t1Range_thr>
    <nbr_above_t2Range_thr>0         </nbr_above_t2Range_thr>
  </Nbr_Spots_Failing_Area_Sizing>
</Rsample>

<Sample>
  <Sample_Type>"Sample"</Sample_Type>
  <Sample_Name>"gel-HM-071"</Sample_Name>
  <Simple_FileName>"gel-HM-071.gif"</Simple_FileName>
  <Sample_Pix_FileName>"demo\ppx\gel-HM-071.gif"</Sample_Pix_FileName>
  <Sample_SSF_FileName>"demo\xml\gel-HM-071-SSF.xml"</Sample_SSF_FileName>
  <cw1>6</cw1>
  <cw2>450</cw2>
  <cw1>68</cw1>
  <cw2>503</cw2>
  <Pix_Height>512</Pix_Height>
  <Pix_Width>512</Pix_Width>
  <PixelSizeMicrons>0.00</PixelSizeMicrons>
  <NbrSpots>3059</NbrSpots>
  <NbrSpotsPrime>1590</NbrSpotsPrime>
  <NbrSpotsOmitted>1469</NbrSpotsOmitted>
  <TotSampleDensity>31534.10</TotSampleDensity>
  <TotSampleDensityPrime>23052.90</TotSampleDensityPrime>
  <TotOmittedDensity>131.30</TotOmittedDensity>
  <TotSampleArea>80117</TotSampleArea>
  <TotSampleAreaPrime>51342</TotSampleAreaPrime>
  <TotSampleAreaOmitted>25905</TotSampleAreaOmitted>
  <PctOmittedToDprimeAcceptedSpots>0.5694 </PctOmittedToDprimeAcceptedSpots>
  <Nbr_Spots_Failing_Area_Sizing>
    <nbr_below_t1Area_thr>2567      </nbr_below_t1Area_thr>
    <nbr_above_t2Area_thr>0          </nbr_above_t2Area_thr>
    <nbr_below_t1Density_thr>4107    </nbr_below_t1Density_thr>
    <nbr_above_t2Density_thr>0       </nbr_above_t2Density_thr>
    <nbr_below_t1Range_thr>0         </nbr_below_t1Range_thr>
    <nbr_above_t2Range_thr>0         </nbr_above_t2Range_thr>
  </Nbr_Spots_Failing_Area_Sizing>
</Sample>
</Pairing_parameters>

<Paired_Spot>
  <LandmarkSet>A</LandmarkSet>
  <R_spotNbr>188</R_spotNbr>
  <R_dxLM>-2</R_dxLM>
  <R_dyLM>-25</R_dyLM>
  <R_xC>207.0</R_xC>
  <R_yC>166.0</R_yC>
  <R_merX1>205</R_merX1>
  <R_merX2>210</R_merX2>
  <R_merY1>163</R_merY1>
  <R_merY2>168</R_merY2>
  <S_spotNbr>759</S_spotNbr>
  <S_dxLM>-2</S_dxLM>
  <S_dyLM>-25</S_dyLM>
  <S_xC>226.0</S_xC>
  <S_yC>166.0</S_yC>
  <S_merX1>224</S_merX1>
  <S_merX2>229</S_merX2>

```

```

<S_merY1>152</S_merY1>
<S_merY2>158</S_merY2>
<PairingCode>P</PairingCode>
<DP>3.0</DP>
<DL>25.1</DL>
<R_area>18</R_area>
<S_area>25</S_area>
<R_dens>1.560</R_dens>
<S_dens>8.352</S_dens>
<R_dPrime>0.826</R_dPrime>
<S_dPrime>3.515</S_dPrime>
<R_volume>1.347</R_volume>
<S_volume>6.484</S_volume>
<R_MaxDens>0.132</R_MaxDens>
<S_MaxDens>0.480</S_MaxDens>
<R_MinDens>0.043</R_MinDens>
<S_MinDens>0.225</S_MinDens>
<R_MeanBkgDens>0.041</R_MeanBkgDens>
<S_MeanBkgDens>0.193</S_MeanBkgDens>
<R_stdDev_X>1.229</R_stdDev_X>
<R_stdDev_Y>1.173</R_stdDev_Y>
<S_stdDev_X>1.292</S_stdDev_X>
<S_stdDev_Y>1.473</S_stdDev_Y>
</Paired_Spot>
<Paired_Spot>
  <LandmarkSet>A</LandmarkSet>
  <R_spotNbr>190</R_spotNbr>
  <R_dxLM>-7</R_dxLM>
  <R_dyLM>-24</R_dyLM>
  <R_xC>202.0</R_xC>
  <R_yC>167.0</R_yC>
  <R_merX1>201</R_merX1>
  <R_merX2>203</R_merX2>
  <R_merY1>164</R_merY1>
  <R_merY2>169</R_merY2>
  <S_spotNbr>759</S_spotNbr>
  <S_dxLM>-2</S_dxLM>
  <S_dyLM>-24</S_dyLM>
  <S_xC>226.0</S_xC>
  <S_yC>167.0</S_yC>
  <S_merX1>224</S_merX1>
  <S_merX2>229</S_merX2>
  <S_merY1>152</S_merY1>
  <S_merY2>158</S_merY2>
  <PairingCode>A</PairingCode>
  <DP>5.4</DP>
  <DL>25.0</DL>
  <R_area>11</R_area>
  <S_area>25</S_area>
  <R_dens>1.470</R_dens>
  <S_dens>8.352</S_dens>
  <R_dPrime>0.961</R_dPrime>
  <S_dPrime>3.515</S_dPrime>
  <R_volume>1.125</R_volume>
  <S_volume>6.484</S_volume>
  <R_MaxDens>0.186</R_MaxDens>
  <S_MaxDens>0.480</S_MaxDens>
  <R_MinDens>0.043</R_MinDens>
  <S_MinDens>0.225</S_MinDens>
  <R_MeanBkgDens>0.046</R_MeanBkgDens>

```

```
<S_MeanBkgDens>0.193</S_MeanBkgDens>
<R_stdDev_X>0.704</R_stdDev_X>
<R_stdDev_Y>1.209</R_stdDev_Y>
<S_stdDev_X>1.292</S_stdDev_X>
<S_stdDev_Y>1.473</S_stdDev_Y>
</Paired_Spot>
```

. . .

```
<Paired_Spot>
  <LandmarkSet>V</LandmarkSet>
  <R_spotNbr>0</R_spotNbr>
  <R_dxLM>0</R_dxLM>
  <R_dyLM>0</R_dyLM>
  <R_xC>0.0</R_xC>
  <R_yC>0.0</R_yC>
  <R_merX1>0</R_merX1>
  <R_merX2>0</R_merX2>
  <R_merY1>0</R_merY1>
  <R_merY2>0</R_merY2>
  <S_spotNbr>1422</S_spotNbr>
  <S_dxLM>159</S_dxLM>
  <S_dyLM>0</S_dyLM>
  <S_xC>446.0</S_xC>
  <S_yC>0.0</S_yC>
  <S_merX1>441</S_merX1>
  <S_merX2>448</S_merX2>
  <S_merY1>493</S_merY1>
  <S_merY2>499</S_merY2>
  <PairingCode>U</PairingCode>
  <DP>166.5</DP>
  <DL>166.5</DL>
  <R_area>0</R_area>
  <S_area>33</S_area>
  <R_dens>0.000</R_dens>
  <S_dens>0.630</S_dens>
  <R_dPrime>0.000</R_dPrime>
  <S_dPrime>0.395</S_dPrime>
  <R_volume>0.000</R_volume>
  <S_volume>0.601</S_volume>
  <R_MaxDens>0.000</R_MaxDens>
  <S_MaxDens>0.041</S_MaxDens>
  <R_MinDens>0.000</R_MinDens>
  <S_MinDens>0.000</S_MinDens>
  <R_MeanBkgDens>0.000</R_MeanBkgDens>
  <S_MeanBkgDens>0.007</S_MeanBkgDens>
  <R_stdDev_X>0.000</R_stdDev_X>
  <R_stdDev_Y>0.000</R_stdDev_Y>
  <S_stdDev_X>1.512</S_stdDev_X>
  <S_stdDev_Y>1.365</S_stdDev_Y>
</Paired_Spot>
```

```
<Global_Spot_pairing_statistics>
  <NbrRsampleSpotsInLMS>738</NbrRsampleSpotsInLMS>
  <NbrSampleSpotsInLMS>1427</NbrSampleSpotsInLMS>
```

```
<Landmark_set_sizes>
  <Landmark>
    <Landmark_name>A</Landmark_name>
    <Landmark_nbr>1</Landmark_nbr>
```

```

    <Nbr_Rsample_spots>12</Nbr_Rsample_spots>
    <Nbr_Sample_spots>18</Nbr_Sample_spots>
</Landmark>
<Landmark>
    <Landmark_name>B</Landmark_name>
    <Landmark_nbr>2</Landmark_nbr>
    <Nbr_Rsample_spots>34</Nbr_Rsample_spots>
    <Nbr_Sample_spots>77</Nbr_Sample_spots>
</Landmark>

. . .

<Landmark>
    <Landmark_name>V</Landmark_name>
    <Landmark_nbr>22</Landmark_nbr>
    <Nbr_Rsample_spots>39</Nbr_Rsample_spots>
    <Nbr_Sample_spots>93</Nbr_Sample_spots>
</Landmark>
</Landmark_set_sizes>

<InitialpairingStats>
    <Nbr_US_spotsPri>608</Nbr_US_spotsPri>
    <Nbr_SP_spotsPri>230</Nbr_SP_spotsPri>
    <Nbr_PP_spotsPri>764</Nbr_PP_spotsPri>
    <Nbr_AP_spotsPri>563</Nbr_AP_spotsPri>
    <Nbr_CP_spotsPri>0</Nbr_CP_spotsPri>
    <Nbr_EP_spotsPri>0</Nbr_EP_spotsPri>
</InitialpairingStats>

<SecondarypairingStats>
    <Nbr_US_spotsSec>594</Nbr_US_spotsSec>
    <Nbr_SP_spotsSec>230</Nbr_SP_spotsSec>
    <Nbr_PP_spotsSec>810</Nbr_PP_spotsSec>
    <Nbr_AP_spotsSec>531</Nbr_AP_spotsSec>
    <Nbr_CP_spotsSec>0</Nbr_CP_spotsSec>
    <Nbr_EP_spotsSec>0</Nbr_EP_spotsSec>
</SecondarypairingStats>
<Primary_SP_PP_pairRate>67.3</Primary_SP_PP_pairRate>
<Secondary_SP_PP_pairRate>70.5</Secondary_SP_PP_pairRate>
<meanDP_SP_PP>10.9</meanDP_SP_PP>
<meanDPprime_SP_PP>17.6</meanDPprime_SP_PP>
</Global_Spot_pairing_statistics>
</SpotList_Pairing>

```

8.4 Examples - part of Sample Pairs File GELLAB-II format

This is part of a Sample Pairs File to illustrate the type of data available as output. This used the default - **spfFormat:X** option.

```

CmpSpots V.0.1.0-pre-Alpha - $Date$ - $Revision$ (Open2Dprot)
Input Rsample file: gel-HM-019.gif
Input Sample file: gel-HM-071.gif
Output Sample Pairs File: demo\xml\gel-HM-071.spf

```

Spot lists parameters and statistics

Rsample Window [14:475,74:509] (pixels) [rows,cols]=[512,512] PixelSizeMicrons=0.00

NbrSpots=3072 NbrSpotsPrime=1717 NbrSpotsOmitted= 1355
 TotSampleDensity=80411.00 TotSampleDensityPrime=26247.40 TotOmittedDensity=121.00
 TotSampleArea=29974 TotSampleAreaPrime=55093 TotSampleAreaOmitted=22282

Sample Window [6:450,68:503] (pixels) [rows,cols]=[512,512] PixelSizeMicrons=0.00
 NbrSpots=1484 NbrSpotsPrime=864 NbrSpotsOmitted= 620
 TotSampleDensity=44242.00 TotSampleDensityPrime=16980.60 TotOmittedDensity=50.80
 TotSampleArea=21961 TotSampleAreaPrime=33078 TotSampleAreaOmitted=11164

Sample pairs threshold sizing parameters

 Threshold Sure-Pair (SP) sizing limit, thrSP: 5.0 (pixels)
 Threshold Possible-Pair (PP) sizing limit, thrPP: 10.0 (pixels)
 Number of alternate landmark sets to check: 1

Switches: -noaccessionFile:accession.xml -noaltLandmarks:1 -nocommutativeLMS
 -nocsamplePrimeData -nodebugBits:0,0 -nodefault -demo
 -nodtd -histGUI -inputFormat:X,X,X -nolandmarkFile:landmark.xml
 -latchLandmarkSpots:6.0 -nopercentDensity -nprojDir:demo\
 -nopropertiesFile:CmpSpots.properties -rsample:gel-HM-019
 -sample:gel-HM-071 -spfFormat:G -secondaryPairing -nossfCW:0,0,0,0
 -thrSP:5.00 -thrPP:10.00 -notimer -nounCorrectedDensity
 -nouupdate:program

```
#A R:209 d(xy)LM1[-2,-26] (xy)C1[205.0,164.0] MER1[202:208,162:167] S:178 d(xy)LM2[-2,-28] (xy)C2
[225.0,148.0] MER2[222:228,147:151]P dP=2.0,dL=28.131area1 22area2 9.190D1gray 5.495D1gray
7.870D'1 1.238D'2 7.756V'1 4.037V'2 0.540Maxd1 0.325Maxd2 0.070MinD1 0.122MinD2 1.348sX1 1.517sX2
1.503sY1 1.155sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.043MnBk1 0.193MnBk2
#A R:197 d(xy)LM1[6,-20] (xy)C1[233.0,156.0] MER1[209:213,164:168] S:219 d(xy)LM2[4,-24] (xy)C2
[211.0,166.0] MER2[230:236,153:160]A dP=4.5,dL=24.343area1 15area2 14.598D1gray 5.336D1gray
7.120D'1 4.780D'2 11.566V'1 3.308V'2 0.540Maxd1 0.376Maxd2 0.174MinD1 0.275MinD2 1.685sX1
1.244sX2 1.792sY1 0.997sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.174MnBk1 0.037MnBk2
#A R:233 d(xy)LM1[-1,-19] (xy)C1[206.0,171.0] MER1[203:208,168:175] S:196 d(xy)LM2[-1,-21] (xy)C2
[226.0,155.0] MER2[224:229,152:158]P dP=2.0,dL=21.034area1 25area2 10.213D1gray 8.352D1gray
8.702D'1 3.515D'2 8.683V'1 6.484V'2 0.480Maxd1 0.480Maxd2 0.122MinD1 0.225MinD2 1.433sX1 1.292sX2
1.779sY1 1.473sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.044MnBk1 0.193MnBk2
#A R:234 d(xy)LM1[6,-18] (xy)C1[213.0,172.0] MER1[209:219,168:176] S:197 d(xy)LM2[6,-20] (xy)C2
[233.0,156.0] MER2[230:236,153:160]P dP=2.0,dL=20.946area1 43area2 18.568D1gray 14.598D1gray
16.779D'1 7.120D'2 17.748V'1 11.566V'2 0.540Maxd1 0.540Maxd2 0.325MinD1 0.174MinD2 2.545sX1
1.685sX2 1.822sY1 1.792sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.039MnBk1 0.174MnBk2
#A R:252 d(xy)LM1[9,-13] (xy)C1[216.0,177.0] MER1[213:219,175:178] S:217 d(xy)LM2[7,-12] (xy)C2
[234.0,164.0] MER2[231:237,161:167]P dP=2.2,dL=15.821area1 37area2 15.298D1gray 9.663D1gray
14.365D'1 0.876D'2 12.723V'1 7.212V'2 1.164Maxd1 0.376Maxd2 0.480MinD1 0.122MinD2 1.570sX1
1.650sX2 0.981sY1 1.639sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.044MnBk1 0.237MnBk2
```

. . . .

```
#A R:278 d(xy)LM1[6,13] (xy)C1[233.0,189.0] MER1[212:221,203:213] S:1111 d(xy)LM2[9,18] (xy)C2
[216.0,208.0] MER2[229:238,185:193]A dP=5.8,dL=20.161area1 64area2 56.456D1gray 15.077D1gray
34.708D'1 10.568D'2 50.597V'1 12.561V'2 1.700Maxd1 0.396Maxd2 0.174MinD1 0.077MinD2 2.089sX1
1.835sX2 2.009sY1 2.441sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.357MnBk1 0.070MnBk2
#A R:223 d(xy)LM1[0,-12] (xy)C1[227.0,164.0] MER1[224:229,160:165] S:223 d(xy)LM2[20,-26] (xy)C2
[227.0,164.0] MER2[224:229,160:165]A dP=5.8,dL=12.016area1 16area2 6.412D1gray 6.412D1gray
2.112D'1 2.112D'2 7.021V'1 7.021V'2 0.540Maxd1 0.540Maxd2 0.225MinD1 0.225MinD2 1.435sX1 1.435sX2
1.278sY1 1.278sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.269MnBk1 0.269MnBk2
#A R:266 d(xy)LM1[23,4] (xy)C1[250.0,180.0] MER1[246:252,177:183] S:266 d(xy)LM2[43,-10] (xy)C2
[250.0,180.0] MER2[246:252,177:183]A dP=7.8,dL=23.338area1 38area2 32.619D1gray 32.619D1gray
21.457D'1 21.457D'2 27.195V'1 27.195V'2 1.392Maxd1 1.392Maxd2 0.480MinD1 0.480MinD2 1.636sX1
1.636sX2 1.685sY1 1.685sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.294MnBk1 0.294MnBk2
#B R:1 d(xy)LM1[0,-82] (xy)C1[208.0,71.0] MER1[204:210,69:73] S:0 d(xy)LM2[31,0] (xy)C2
```

```

[0.0,0.0] MER2[0:0,0:0]U dL=10.0,dL=87.324area1 0area2 0.971Dlgray 0.000Dlgray 0.971D'1 0.000D'2
0.995V'1 0.000V'2 0.070Maxd1 0.000Maxd2 0.027MinD1 0.000MinD2 1.572sX1 0.000sX2 1.283sY1 0.000sY2
0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.000MnBk1 0.000MnBk2
#B R:26 d(xy)LM1[0,-77] (xy)C1[158.0,76.0] MER1[155:162,73:78] S:0 d(xy)LM2[-17,0] (xy)C2
[0.0,0.0] MER2[0:0,0:0]U dL=10.4,dL=79.634area1 0area2 14.270Dlgray 0.000Dlgray 13.703D'1
0.000D'2 11.852V'1 0.000V'2 0.915Maxd1 0.000Maxd2 0.027MinD1 0.000MinD2 1.840sX1 0.000sX2
0.993sY1 0.000sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.017MnBk1 0.000MnBk2
#B R:27 d(xy)LM1[0,-77] (xy)C1[165.0,76.0] MER1[163:167,73:78] S:0 d(xy)LM2[-17,0] (xy)C2
[0.0,0.0] MER2[0:0,0:0]U dL=10.8,dL=78.119area1 0area2 6.880Dlgray 0.000Dlgray 6.598D'1 0.000D'2
5.253V'1 0.000V'2 0.667Maxd1 0.000Maxd2 0.027MinD1 0.000MinD2 1.187sX1 0.000sX2 0.936sY1 0.000sY2
0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.015MnBk1 0.000MnBk2
#B R:28 d(xy)LM1[0,-77] (xy)C1[170.0,76.0] MER1[168:171,73:78] S:0 d(xy)LM2[-17,0] (xy)C2
[0.0,0.0] MER2[0:0,0:0]U dL=13.5,dL=77.416area1 0area2 2.850Dlgray 0.000Dlgray 2.613D'1 0.000D'2
2.701V'1 0.000V'2 0.428Maxd1 0.000Maxd2 0.027MinD1 0.000MinD2 0.925sX1 0.000sX2 0.962sY1 0.000sY2
0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.015MnBk1 0.000MnBk2

. . .

#U R:0 d(xy)LM1[0,0] (xy)C1[0.0,0.0] MER1[0:0,0:0] S:862 d(xy)LM2[20,65] (xy)C2[256.0,498.0]
MER2[252:259,494:501]U dL=68.0,dL=68.00area1 52area2 0.000Dlgray 36.111Dlgray 34.904D'1 34.904D'2
30.105V'1 30.105V'2 0.000Maxd1 1.392Maxd2 0.000MinD1 0.070MinD2 0.000sX1 1.777sX2 0.000sY1
1.716sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.000MnBk1 0.023MnBk2
#U R:0 d(xy)LM1[0,0] (xy)C1[0.0,0.0] MER1[0:0,0:0] S:863 d(xy)LM2[-68,67] (xy)C2[168.0,500.0]
MER2[164:171,497:502]U dL=95.5,dL=95.50area1 38area2 0.000Dlgray 10.867Dlgray 9.917D'1 9.917D'2
9.453V'1 9.453V'2 0.000Maxd1 0.540Maxd2 0.000MinD1 0.070MinD2 0.000sX1 1.789sX2 0.000sY1 1.380sY2
0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.000MnBk1 0.025MnBk2
#U R:0 d(xy)LM1[0,0] (xy)C1[0.0,0.0] MER1[0:0,0:0] S:864 d(xy)LM2[-86,68] (xy)C2[150.0,501.0]
MER2[147:153,499:502]U dL=99.0,dL=109.60area1 19area2 0.000Dlgray 0.975Dlgray 0.534D'1 0.534D'2
1.143V'1 1.143V'2 0.000Maxd1 0.122Maxd2 0.000MinD1 0.027MinD2 0.000sX1 1.457sX2 0.000sY1 0.909sY2
0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.000MnBk1 0.023MnBk2
#V R:1510 d(xy)LM1[-46,13] (xy)C1[241.0,431.0] MER1[238:244,427:435] S:757 d(xy)LM2[-40,8] (xy)C2
[276.0,400.0] MER2[272:280,396:404]P dP=7.8,dL=47.846area1 50area2 11.915Dlgray 36.901Dlgray
10.297D'1 34.401D'2 9.979V'1 32.682V'2 0.500Maxd1 1.392Maxd2 0.043MinD1 0.174MinD2 1.649sX1
1.809sX2 1.707sY1 1.830sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.035MnBk1 0.050MnBk2
#V R:658 d(xy)LM1[0,-63] (xy)C1[322.0,355.0] MER1[318:326,351:360] S:0 d(xy)LM2[29,0] (xy)C2
[0.0,0.0] MER2[0:0,0:0]U dL=13.4,dL=72.157area1 0area2 42.779Dlgray 0.000Dlgray 41.829D'1
0.000D'2 38.131V'1 0.000V'2 1.700Maxd1 0.000Maxd2 0.070MinD1 0.000MinD2 1.887sX1 0.000sX2
1.677sY1 0.000sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.017MnBk1 0.000MnBk2
#V R:671 d(xy)LM1[0,-56] (xy)C1[371.0,362.0] MER1[368:376,359:365] S:0 d(xy)LM2[55,0] (xy)C2
[0.0,0.0] MER2[0:0,0:0]U dL=38.9,dL=101.043area1 0area2 1.887Dlgray 0.000Dlgray 1.569D'1
0.000D'2 2.180V'1 0.000V'2 0.122Maxd1 0.000Maxd2 0.000MinD1 0.000MinD2 1.740sX1 0.000sX2 1.452sY1
0.000sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.007MnBk1 0.000MnBk2
#V R:660 d(xy)LM1[37,-47] (xy)C1[353.0,345.0] MER1[326:334,359:366] S:674 d(xy)LM2[43,-55] (xy)C2
[330.0,363.0] MER2[351:356,341:349]A dP=10.0,dL=69.828area1 49area2 1.416Dlgray 20.034Dlgray
1.016D'1 19.490D'2 1.828V'1 16.801V'2 0.122Maxd1 0.723Maxd2 0.027MinD1 0.122MinD2 1.398sX1
1.922sX2 1.514sY1 1.705sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.014MnBk1 0.011MnBk2

. . .

#V R:0 d(xy)LM1[0,0] (xy)C1[0.0,0.0] MER1[0:0,0:0] S:827 d(xy)LM2[65,60] (xy)C2[381.0,452.0]
MER2[378:384,449:455]U dL=99.0,dL=101.40area1 27area2 0.000Dlgray 0.484Dlgray 0.339D'1 0.339D'2
0.474V'1 0.474V'2 0.000Maxd1 0.041Maxd2 0.000MinD1 0.000MinD2 0.000sX1 1.437sX2 0.000sY1 1.134sY2
0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.000MnBk1 0.005MnBk2
#V R:0 d(xy)LM1[0,0] (xy)C1[0.0,0.0] MER1[0:0,0:0] S:836 d(xy)LM2[5,70] (xy)C2[321.0,462.0] MER2
[317:324,459:465]U dL=70.2,dL=70.20area1 40area2 0.000Dlgray 3.944Dlgray 3.444D'1 3.444D'2
3.371V'1 3.371V'2 0.000Maxd1 0.174Maxd2 0.000MinD1 0.041MinD2 0.000sX1 1.943sX2 0.000sY1
1.408sY2 0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.000MnBk1 0.013MnBk2
#V R:0 d(xy)LM1[0,0] (xy)C1[0.0,0.0] MER1[0:0,0:0] S:855 d(xy)LM2[15,96] (xy)C2[331.0,488.0]
MER2[327:333,486:491]U dL=97.2,dL=97.20area1 22area2 0.000Dlgray 0.580Dlgray 0.462D'1 0.462D'2
0.480V'1 0.480V'2 0.000Maxd1 0.041Maxd2 0.000MinD1 0.000MinD2 0.000sX1 1.330sX2 0.000sY1 1.240sY2

```

```

0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.000MnBk1 0.005MnBk2
#V R:0 d(xy)LM1[0,0] (xy)C1[0.0,0.0] MER1[0:0,0:0] S:857 d(xy)LM2[37,103] (xy)C2[353.0,495.0]
MER2[349:357,491:498]U dL=99.0,dL=109.40area1 47area2 0.000Dlgray 1.691Dlgray 1.439D'1 1.439D'2
1.617V'1 1.617V'2 0.000Maxd1 0.070Maxd2 0.000MinD1 0.000MinD2 0.000sX1 1.849sX2 0.000sY1 1.774sY2
0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.000MnBk1 0.005MnBk2
#V R:0 d(xy)LM1[0,0] (xy)C1[0.0,0.0] MER1[0:0,0:0] S:859 d(xy)LM2[130,104] (xy)C2[446.0,496.0]
MER2[441:448,493:499]U dL=99.0,dL=177.10area1 33area2 0.000Dlgray 0.630Dlgray 0.395D'1 0.395D'2
0.601V'1 0.601V'2 0.000Maxd1 0.041Maxd2 0.000MinD1 0.000MinD2 0.000sX1 1.512sX2 0.000sY1 1.365sY2
0.00sdMX,0.00sdMY,0.00sdMD0.00sdMA,0#G,0.000MnBk1 0.007MnBk2

```

Summary: of paired-spot statistics

Rsample has 1564, Sample 864 spots in all landmark sets.

After Initial pairing:

```

US 107
SP 312
PP 882
AP 671
CP 0
EP 0

```

$0.5(SP+PP)/(|G1| \text{ MIN } |G2|)=69.10\%$

After secondary pairing:

```

US 95
SP 312
PP 918
AP 647
CP 0
EP 0

```

After secondary pairing: $0.5(SP+PP)/(|G1| \text{ MIN } |G2|)=71.2\%$

mean dP(SP+PP)=1.88, mean dP'((|G1|+|G2|)/(SP+PP))=0.63

List of image files and generated files

```

Input Rsample pix file [demo\ppx\gel-HM-019.gif]
Input Sample pix file [demo\ppx\gel-HM-071.gif]
Input Rsample SSF file [demo\xml\gel-HM-019.xml]
Input Sample SSF file [demo\xml\gel-HM-071.xml]
Output SPF file [demo\xml\gel-HM-071.spf]

```

FINISHED! The Sample Paired-spotlist File (SPF), is demo\xml\gel-HM-071.spf

Run time =0:0:1 (H:M:S) or 1.6 seconds

8.5 Examples - using command line processing for batch

It is possible to run the CmpSpots from the command line in your operating system (substituting "-nogui" for the "-gui" switch). We give two examples doing this. The first example shows a script for the Microsoft Windows batch (.bat) file [demo-CmpSpots.bat](#) file (available on the [Files Mirror](#)). The [second example](#) shows the same commands in a shell script for a Unix operating system (Linux, MacOS, Solaris, etc.) that you can download with [demo-CmpSpots.sh](#).

Command line scripts to run the GUI version on one sample are in the [second example](#) and [forth example](#). These may be downloaded from [demo-CmpSpots-GUI.bat](#) and [demo-CmpSpots-GUI.sh](#).

The data for these scripts is in the **demo/** directory available with the installation. Alternatively, you can download the demo data from the Files Mirror as [Demo.zip](#).

The installation packages are available from the [Files mirror](#) under the CmpSpots releases.

8.5.1 Examples - batch processing under Microsoft Windows

```
REM File: demo-CmpSpots.bat - pair spots to the Reference sample for a list of
samples
REM This example assumes that all .jar files listed below and demo/ directory are
REM in the current directory. Modify for other situations.
REM
REM The JDK should be installed and version 1.4 or later is required.
REM You can download the latest JDK from http://java.sun.com/
REM
REM The files needed are listed below:
REM JAR files required and mentioned in manifest:
REM JAR files:
REM      xml-apis.jar xercesImpl.jar jai_codec.jar jai_core.jar O2Plib.jar
REM
REM demo image files:
REM      demo/ppx/gel-HM-019 (Reference sample)
REM      demo/ppx/gel-HM-071
REM      demo/ppx/gel-HM-087
REM      demo/ppx/gel-HM-096
REM Accession database file is in:
REM      demo/xml/accession.xml
REM Landmark database file is in:
REM      demo/xml/landmark.xml
REM Sample Spot-list files (SSF) are in:
REM      demo/xml/gel-HM-019-SSF.xml
REM      demo/xml/gel-HM-071-SSF.xml
REM      demo/xml/gel-HM-087-SSF.xml
REM      demo/xml/gel-HM-096-SSF.xml
REM Generated Sample Paired-spot-list Files (SPF) are saved in:
REM      demo/xml/
REM Generated images are saved in:
REM      demo/tmp/
REM
REM P. Lemkin $Date$

echo "demo-CmpSpots.bat"

pwd
date /T

java -Xmx256M -jar CmpSpots.jar -default -nogui -project:demo\ -sample:gel-HM-
```

```

071 -rsample:gel-HM-019
java -Xmx256M -jar CmpSpots.jar -default -nogui -project:demo\ -sample:gel-HM-
087 -rsample:gel-HM-019
java -Xmx256M -jar CmpSpots.jar -default -nogui -project:demo\ -sample:gel-HM-
096 -rsample:gel-HM-019

echo "-- Finished pairing the samples ---"

date /T

```

8.5.1.1 Examples - batch processing with GUI under Microsoft Windows

```

REM File: demo-CmpSpots-GUI.bat - pair spots to the Reference sample for a
sample using the GUI
REM This example assumes that all .jar files listed below and demo/ directory are
REM in the current directory. Modify for other situations.
REM
REM The JDK should be installed and version 1.4 or later is required.
REM You can download the latest JDK from http://java.sun.com/
REM
REM The files needed are listed below:
REM JAR files required and mentioned in manifest:
REM      xml-apis.jar xercesImpl.jar jai_codec.jar jai_core.jar O2Plib.jar
REM
REM demo image files:
REM      demo/ppx/gel-HM-019 (Reference sample)
REM      demo/ppx/gel-HM-071
REM      demo/ppx/gel-HM-087
REM      demo/ppx/gel-HM-096
REM Accession database file is in:
REM      demo/xml/accession.xml
REM Landmark database file is in:
REM      demo/xml/landmark.xml
REM Sample Spot-list files (SSF) are in:
REM      demo/xml/gel-HM-019-SSF.xml
REM      demo/xml/gel-HM-071-SSF.xml
REM      demo/xml/gel-HM-087-SSF.xml
REM      demo/xml/gel-HM-096-SSF.xml
REM Generated Sample Paired-spot-list Files (SPF) are saved in::
REM      demo/xml/
REM Generated images are saved in:
REM      demo/tmp/
REM
REM P. Lemkin $Date$

echo "demo-CmpSpots-GUI.bat"

```

```

pwd
date /T

java -Xmx256M -jar CmpSpots.jar -demo -gui -project:demo\ -sample:gel-HM-071 -
rsample:gel-HM-019

echo "-- Finished pairing the sample ---"

date /T

```

8.5.2 Examples - batch processing under Unix

Because java is relatively operating system independent, the same java command lines are used with the "\" changed to "/" from Windows to Unix file path convention.

```

#!/bin/sh
# File: demo-CmpSpots.sh - Unix script to pair spots to the
# Reference sample for a list of samples
# This example assumes that all .jar files listed below and demo/ directory are
# in the current directory. Modify for other situations.
#
# The JDK should be installed and version 1.4 or later is required.
# You can download the latest JDK from http://java.sun.com/
#
# The files needed are listed below:
# JAR files required and mentioned in manifest:
#     xml-apis.jar xercesImpl.jar jai_codec.jar jai_core.jar O2Plib.jar
#
# demo sample image files:
#     demo/ppx/gel-HM-019 (Reference sample)
#     demo/ppx/gel-HM-071
#     demo/ppx/gel-HM-087
#     demo/ppx/gel-HM-096
# Accession database file is in:
#     demo/xml/accession.xml
# Landmark database file is in:
#     demo/xml/landmark.xml
# Sample Spot-list files (SSF) are in:
#     demo/xml/gel-HM-019-SSF.xml
#     demo/xml/gel-HM-071-SSF.xml
#     demo/xml/gel-HM-087-SSF.xml
#     demo/xml/gel-HM-096-SSF.xml:
# Generated Sample Paired-spot-list Files (SPF) are saved in:
#     demo/xml/
# Generated images are saved in:
#     demo/tmp/
#

```



```
# P. Lemkin $Date$

echo "demo-CmpSpots.sh"

pwd
date

java -Xmx256M -jar CmpSpots.jar -default -nogui -project:demo/ -sample:gel-HM-071 -rsample:gel-HM-019
java -Xmx256M -jar CmpSpots.jar -default -nogui -project:demo/ -sample:gel-HM-087 -rsample:gel-HM-019
java -Xmx256M -jar CmpSpots.jar -default -nogui -project:demo/ -sample:gel-HM-096 -rsample:gel-HM-019

echo "-- Finished pairing the samples ---"

date
```

8.5.2.1 Examples - batch processing with GUI under Unix

```
#!/bin/sh
# File: demo-CmpSpots-GUI.sh - Unix script to pair spots to the
# Reference sample for a sample using the GUI.
# This example assumes that all .jar files listed below and demo/ directory are
# in the current directory. Modify for other situations.
#
# The JDK should be installed and version 1.4 or later is required.
# You can download the latest JDK from http://java.sun.com/
#
# The files needed are listed below:
# JAR files required and mentioned in manifest:
#     xml-apis.jar xercesImpl.jar jai_codec.jar jai_core.jar O2Plib.jar
#
# demo sample image files:
#     demo/ppx/gel-HM-019 (Reference sample)
#     demo/ppx/gel-HM-071
# Accession database file is in:
#     demo/xml/accession.xml
# Landmark database file is in:
#     demo/xml/landmark.xml
# Sample Spot-list files (SSF) are in:
#     demo/xml/gel-HM-019-SSF.xml
#     demo/xml/gel-HM-071-SSF.xml
#     demo/xml/gel-HM-087-SSF.xml
#     demo/xml/gel-HM-096-SSF.xml:
# Generated Sample Paired-spot-list Files (SPF) are saved in:
#     demo/xml/
```

```
# Generated images are saved in:
#     demo/tmp/
#
# P. Lemkin $Date$

echo "demo-CmpSpots-GUI.sh"

pwd
date

java -Xmx256M -jar CmpSpots.jar -demo -gui -project:demo/ -sample:gel-HM-071 -
rsample:gel-HM-019

echo "-- Finished pairing the samples ---"

date
```

9. CmpSpots References

These papers (a subset of the [GELLAB-II](#) papers), reference the GELLAB-II spot pairing program. The Open2Dprot Java-language CmpSpots program was derived from the GELLAB-II C-language program as well as from code from the [MAExplorer](#) and [Flicker](#) projects. New Java code was added as well. Although CmpSpots has been enhanced in many ways, the basic algorithm is similar so these papers may be useful for more details on the algorithm.

1. Lipkin, L.E., Lemkin, P.F. (1980) Database techniques for multiple PAGE (2D gel) analysis. *Clinical Chemistry* **26**, 1403-1413.
2. Lemkin, P., Lipkin, L. (1981) GELLAB: A computer system for 2D gel electrophoresis analysis. I. Segmentation and preliminaries. *Computers in Biomedical Research* **14**, 272-297.
3. Lemkin, P., Lipkin, L. (1981) GELLAB: A computer system for 2D gel electrophoresis analysis. II. Pairing Spots. *Computers in Biomedical Research* **14**, 355-380.
4. Lester, E.P., Lemkin, P.F., Lipkin, L.E. (1981) New Dimensions in Protein Analysis - 2D Gels Coming of Age Through Image Processing, Invited paper, *Analytical Chemistry* **53**, 390A-397A.
5. Lemkin, P.F., Lipkin, L.E. (1983) Database Techniques for 2D Electrophoretic Gel Analysis, in Computing in Biological Science, Elsevier/North-Holland, M. Geisow, A. Barrett (eds), 181-226.
6. Lemkin, P.F., Lipkin, L.E., Lester, E.P. (1982) Extensions to the GELLAB 2D Electrophoresis Gel Analysis System. Paper given at "Clinical Applications of 2D Electrophoresis", Mayo Clinic, Nov. 15-18, 1981. *Clinical Chemistry* 840-849.
7. Lemkin, P.F., Lester, E.P. (1989) Database and Search Techniques for 2D Gel Protein Data: A

Comparison of Paradigms For Exploratory Data Analysis and Prospects for Biological Modeling, *Electrophoresis* **10**(2): 122-140.

8. Lemkin, P.F. (1989) GELLAB-II, A workstation based 2D electrophoresis gel analysis system, in proceedings of Two-Dimensional Electrophoresis T. Endler, S.Hanash (Eds), Vienna Austria, Nov 8-11, 1988, VCH Press, W.Germany. pp 53-57.
9. Lemkin, P.F., Rogan, P. (1991) Automatic Detection of noisy spots in two-dimensional Southern Blots, *Applied and Theoretical Electrophoresis* **2**, 141-149.
10. Myrick, J.E., Lemkin, P.F., Robinson, M.K., Upton, K.M. (1993) Comparison of the BioImage Visage (TM) 2,000 and the GELLAB-II two-dimensional electrophoretic analysis systems. *Applied and Theoretical Electrophoresis* **3**, 335-346.

CmpSpots is a contributed program available at [open2dprot.](http://open2dprot.sourceforge.net/CmpSpots)

sourceforge.net/CmpSpots

Revised: 05/31/2006

Powered by

SourceForge
Logo

Sample Screen Shots for CmpSpots 2D Spot Pairing - pre-alpha version

To give the flavor of running the spot pairing program, we provide a few screen shots of the graphical user interface and some Image Viewer images generated by the program.

[1. Initial user interface window](#)

[2. Initial command line options tool](#)

[3. Using Image Viewer to investigate paired results](#)

- 3.1 [Image viewer - no overlay and no filters](#)
- 3.2 [Image viewer - landmarks](#)
- 3.3 [Image viewer - landmark sets radii](#)
- 3.4 [Image viewer - radii labels no filter](#)
- 3.5 [Image viewer - labels no filter](#)
- 3.6 [Image viewer - labels filter LM set P](#)
- 3.7 [Image viewer - radii labels filter LM set P](#)
- 3.8 [Image viewer - vector labels no filter](#)
- 3.9 [Image viewer - vector labels filter LM set V](#)
- 3.10 [Image viewer - vector labels radius filter LM set V](#)

- 3.11 [Image viewer - labels filter LM sets A, B, C, D](#)
- 3.12 [Image viewer - histogram filter](#)
- 3.13 [Image viewer - histogram filter by dP limit](#)
- 3.14 [Image viewer - histogram filter dL limit](#)

1. Initial Report Window



Clear

SaveAs

Stop spot pairing

Edit options

Pair spots

Image viewer

Close

2. Initial command line options window

Enter sample image and select switch options

Change the startup parameters, then press 'Set new options' button to save them.
At that point, you can press 'Pair spots' to pair the Rsample and Sample with the new parameters. You may also enter new Rsample and Sample data file names and/or edit switch options and threshold sliders.

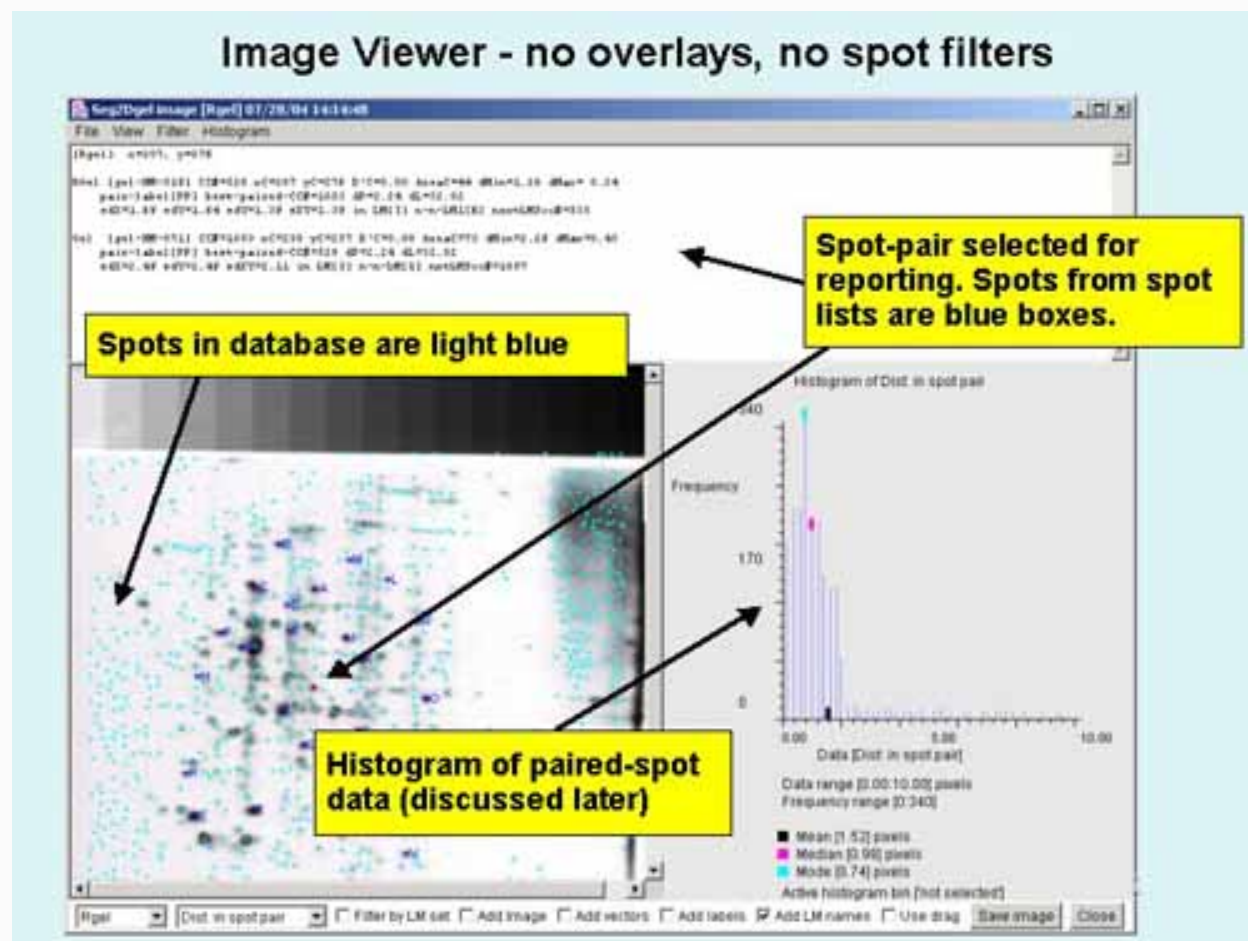
<input type="checkbox"/> -accessionFile	accession.xml	
<input type="checkbox"/> -allLandmarks	1	
<input type="checkbox"/> -commutativeLMS	---	Number alternate landmarks: 1
<input type="checkbox"/> -cSamplePrimeData	---	
<input type="checkbox"/> -debugBits	0.0	
<input type="checkbox"/> -default	---	
<input checked="" type="checkbox"/> -demo	---	
<input type="checkbox"/> -dtd	---	
<input checked="" type="checkbox"/> -histGUI	---	
<input checked="" type="checkbox"/> -inputFormat	X,X,X	Max latch landmark threshold (pixels): 6.0
<input type="checkbox"/> -landmarkFile	landmark.xml	
<input checked="" type="checkbox"/> -latchLandmarkSpots	6.0	
<input type="checkbox"/> -percentDensity	---	
<input type="checkbox"/> -projDir Project directory	demo\	
<input type="checkbox"/> -propertiesFile Properties file	CmpSpots.properties	
<input checked="" type="checkbox"/> -rsample Rsample file	gel-HM-019	Max SP distance threshold (pixels): 5.00
<input checked="" type="checkbox"/> -sample Sample file	gel-HM-071	
<input checked="" type="checkbox"/> -spFormat X	X	
<input checked="" type="checkbox"/> -secondaryPairing	---	
<input type="checkbox"/> -ssfCW	0.0,0.0	
<input checked="" type="checkbox"/> -thrSP	5.00	
<input checked="" type="checkbox"/> -thrPP	10.00	Max PP distance threshold (pixels): 10.00
<input type="checkbox"/> -timer	---	
<input type="checkbox"/> -unCorrectedDensity	---	
<input type="checkbox"/> -update program	program	

Reset defaults Set new default Cancel

Using Image Viewer to investigate paired results

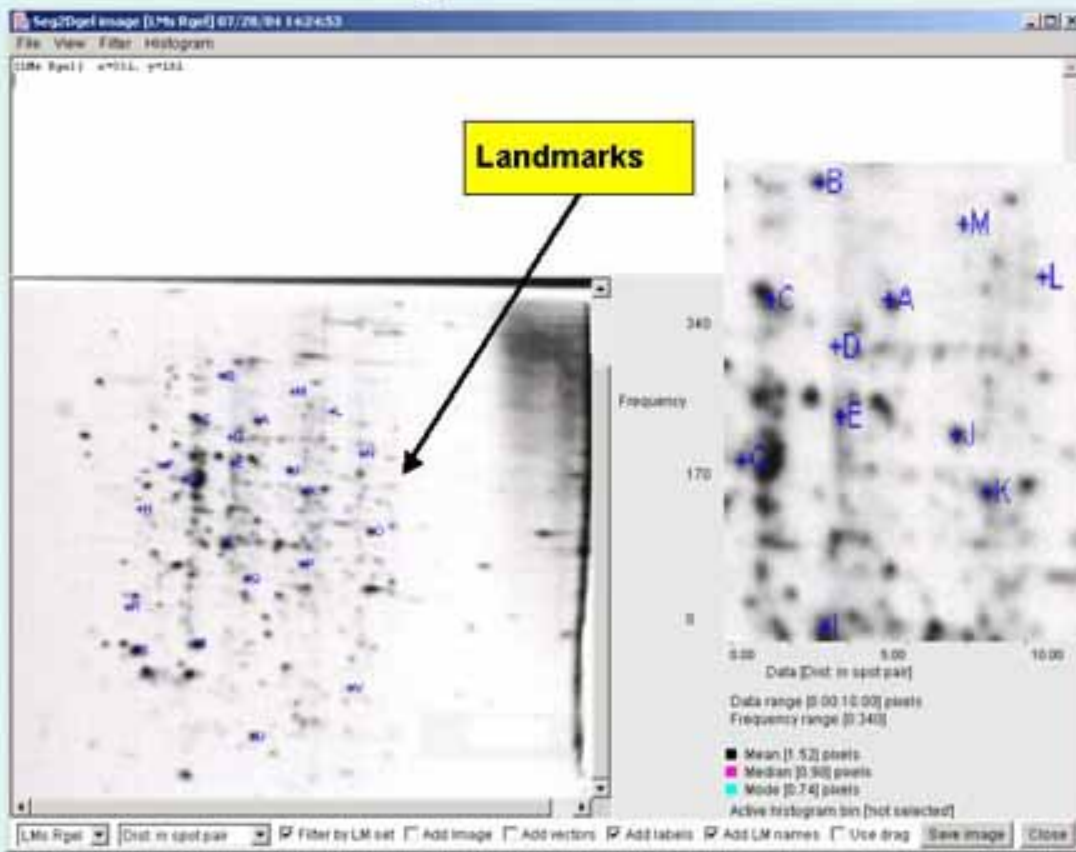
If you enable the CmpSpots to generate images, you can view the pairing results using the ImageViewer. This is enabled using the **-gui** switch.

3.1 Image viewer - no overlay and no filters



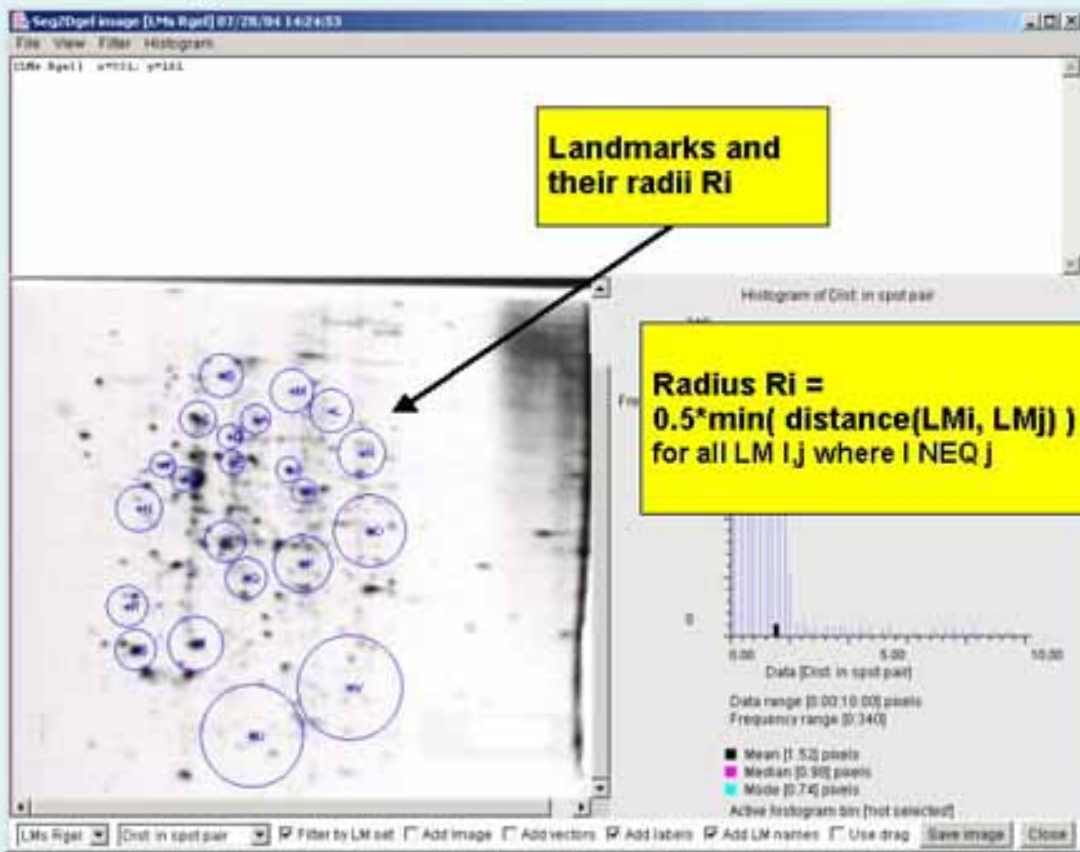
3.2 Image viewer - landmarks

Image Viewer - landmarks



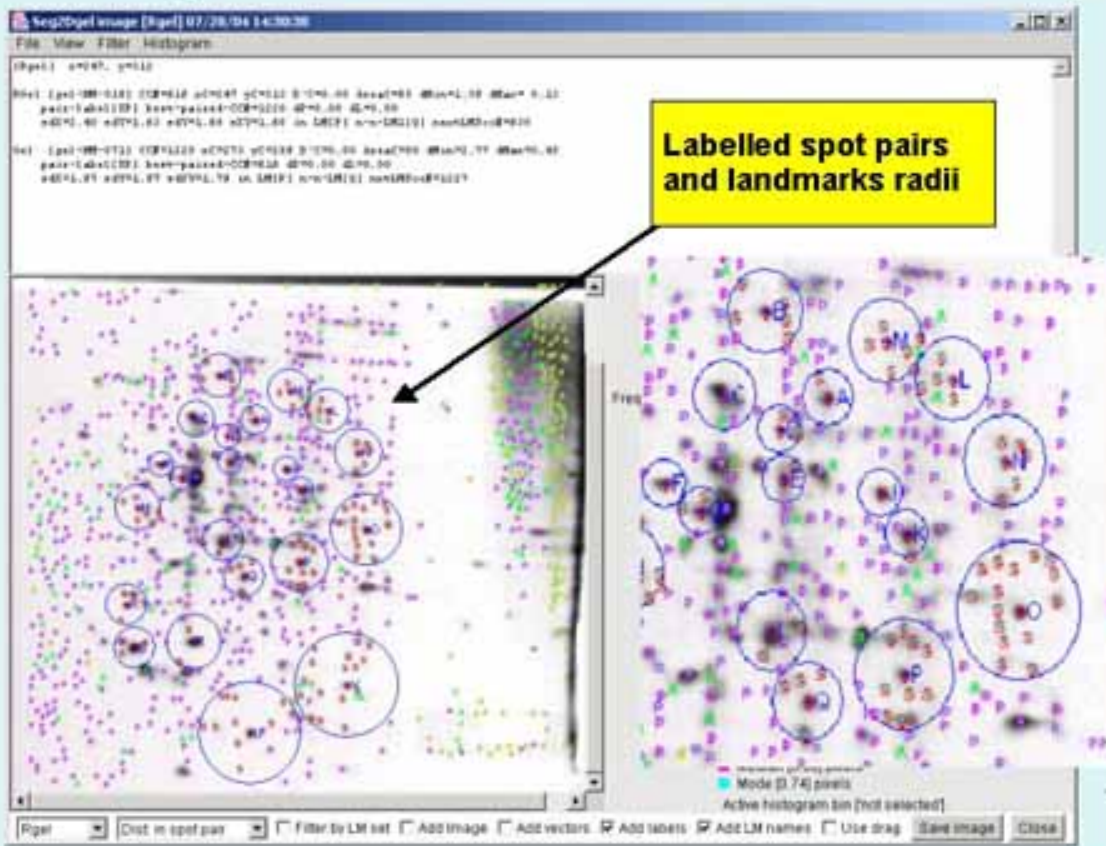
3.3 Image viewer - landmark sets radii

Image Viewer - landmarks and their effective radii

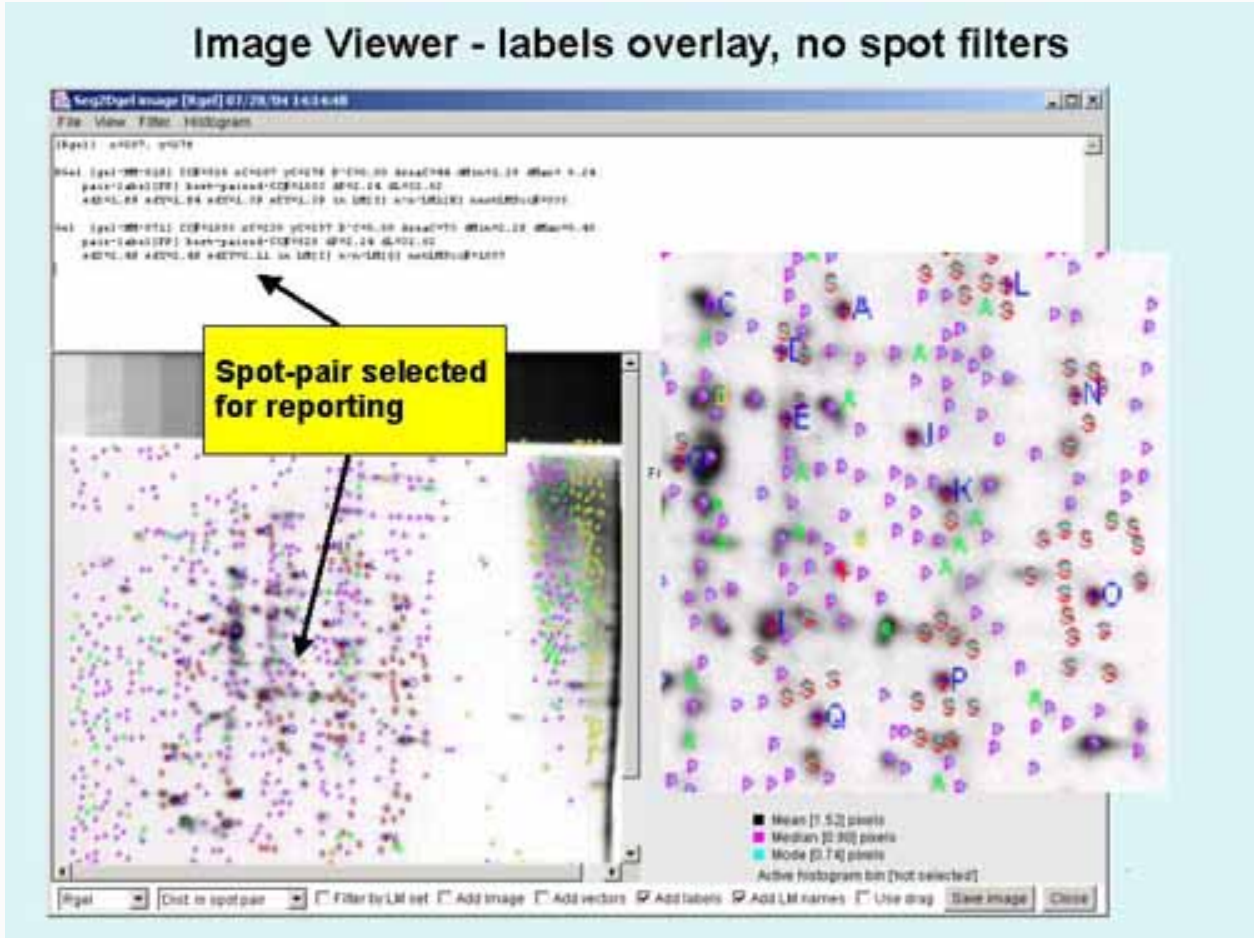


3.4 Image viewer - radii labels no filter

Image Viewer - labels and landmarks radii overlays

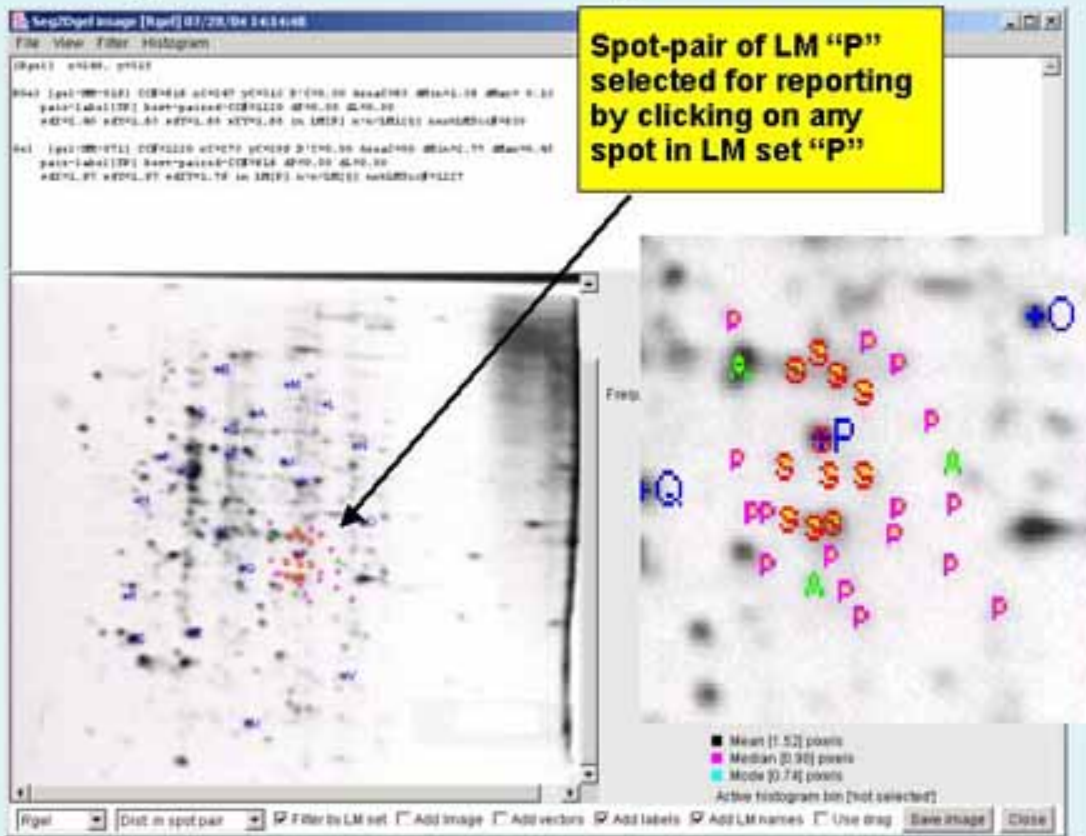


3.5 Image viewer - labels no filter



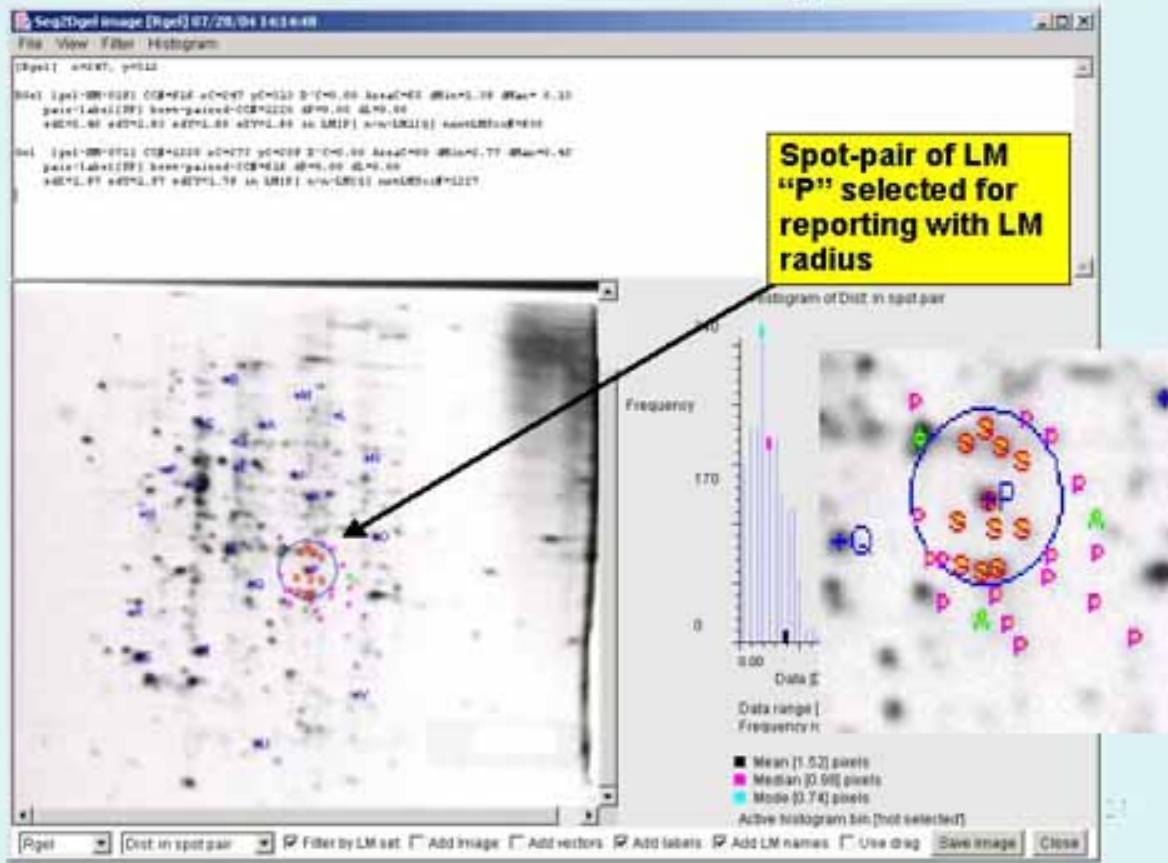
3.6 Image viewer - labels filter LM set P

Image Viewer - labels overlay, current landmark set "P"

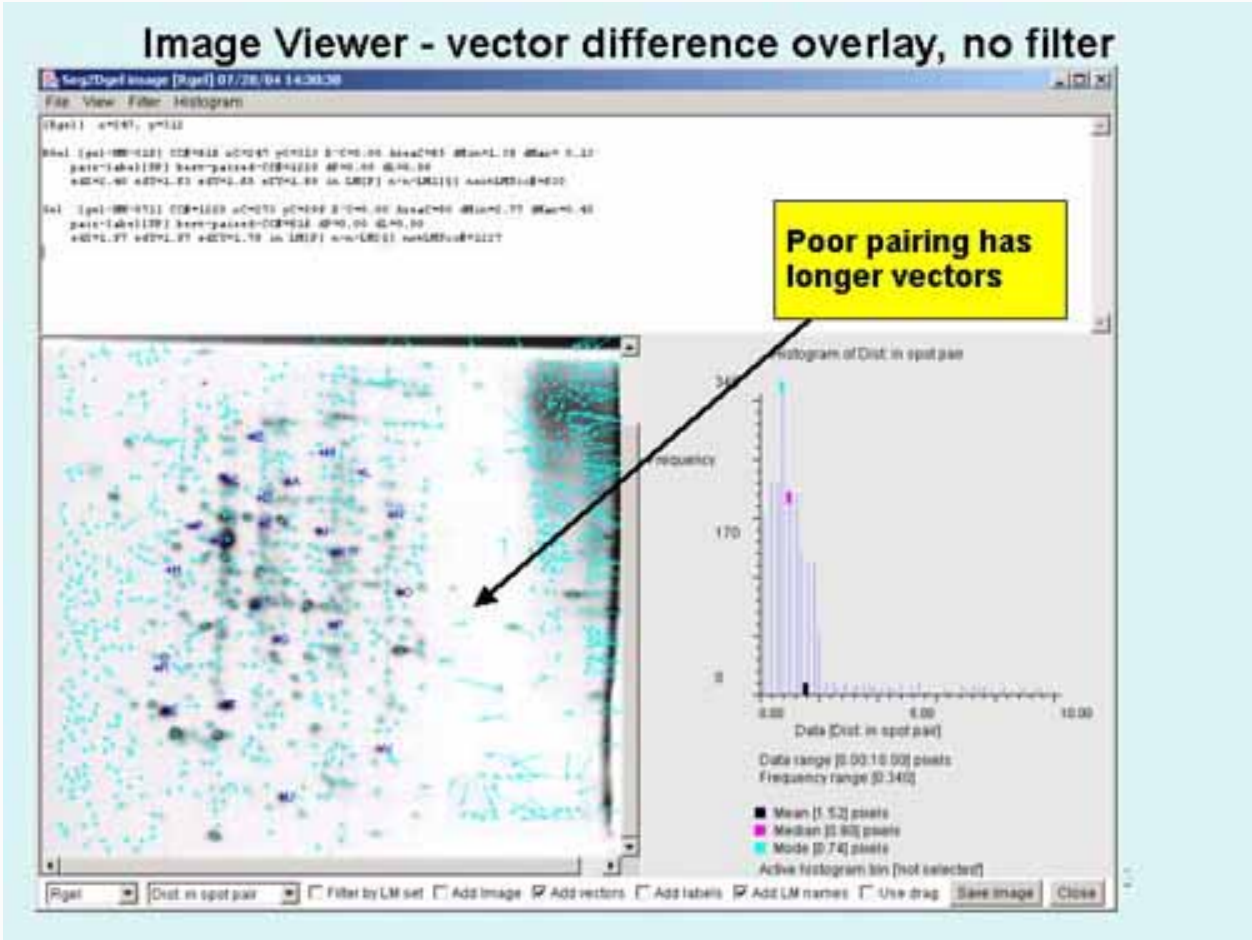


3.7 Image viewer - radii labels filter LM set P

Image Viewer - labels & radius overlay, landmark set "P"

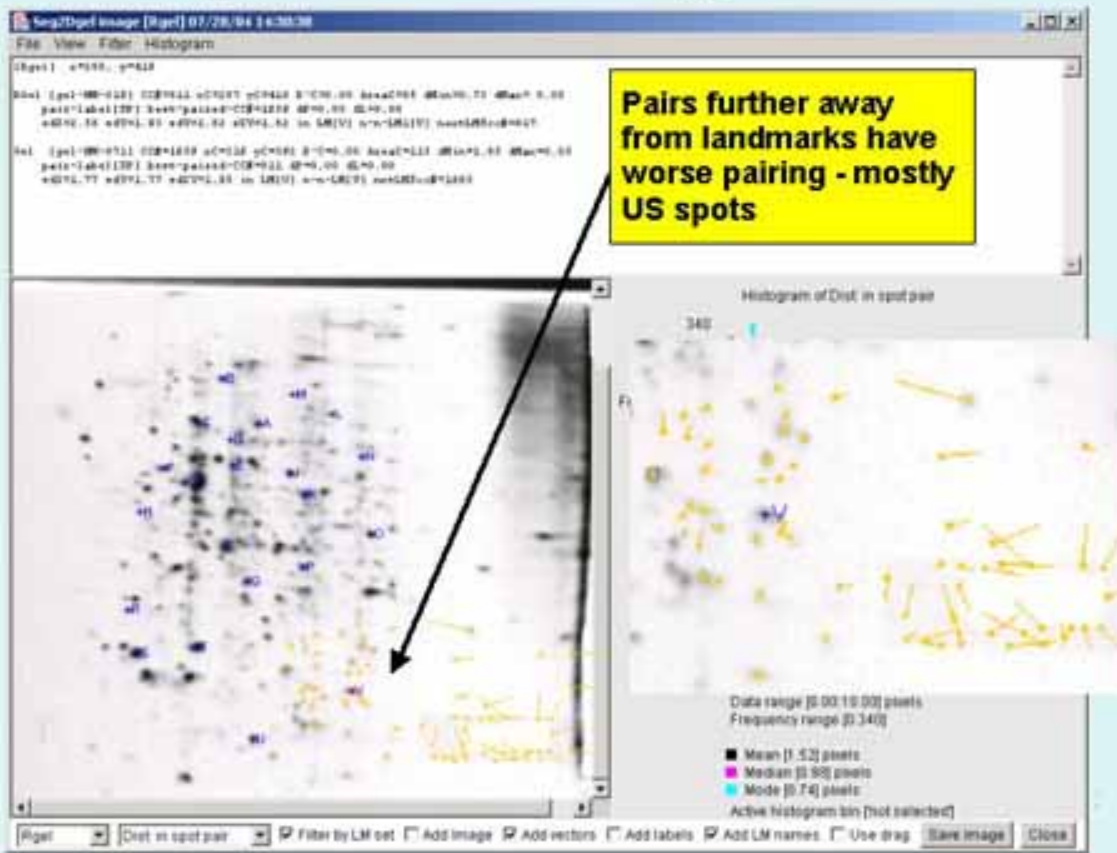


3.8 Image viewer - vector labels no filter



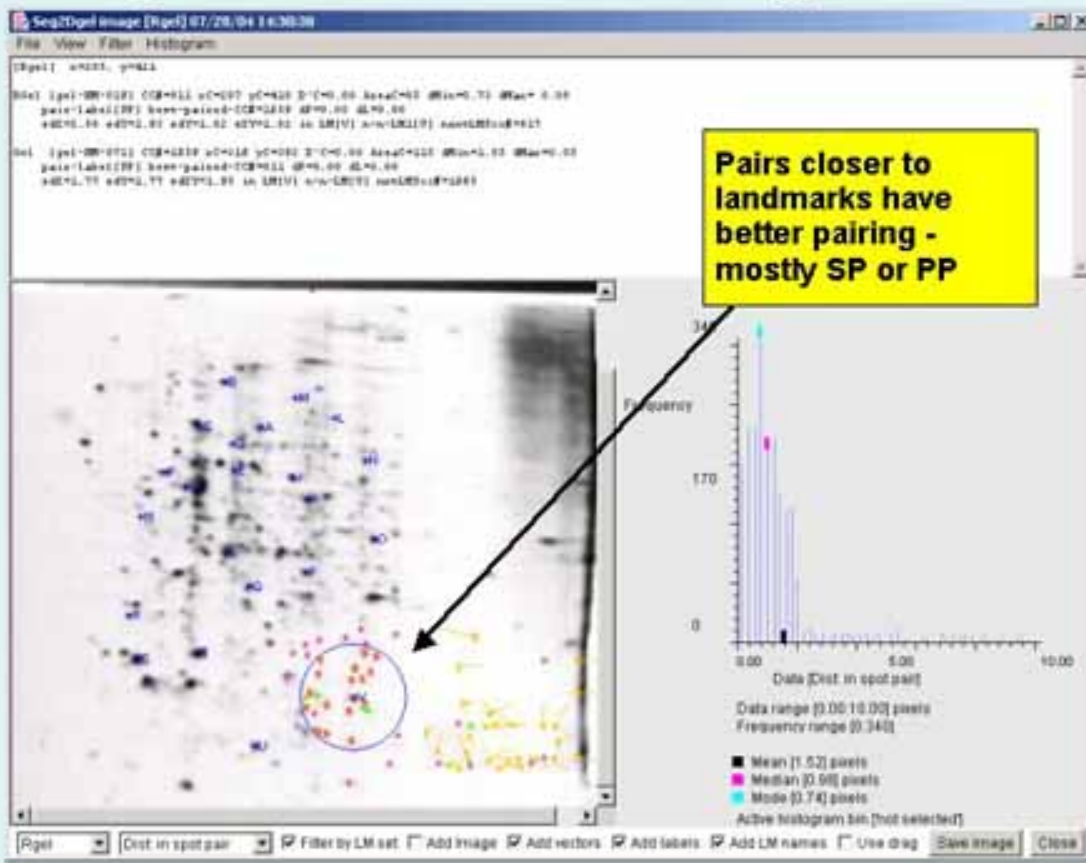
3.9 Image viewer - vector labels filter LM set V

Image Viewer - vector overlay, landmark set "V"



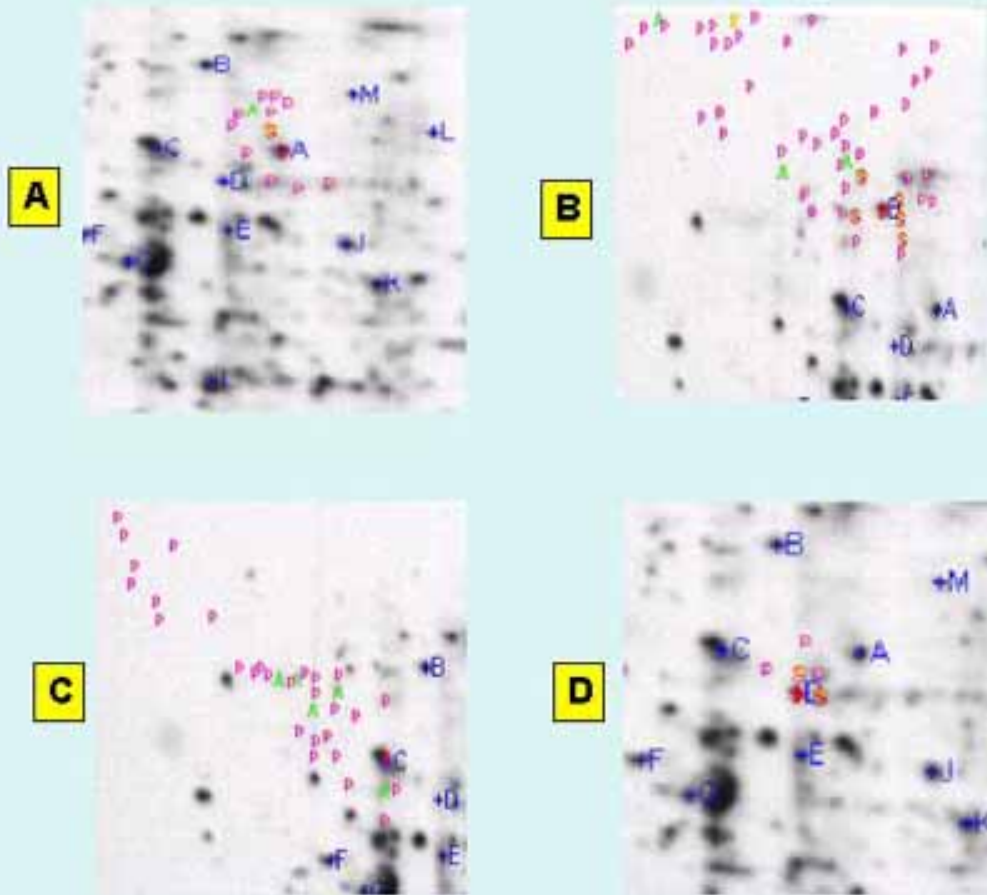
3.10 Image viewer - vector labels radius filter LM set V

Image Viewer - labels & radius overlay, landmark set "V"



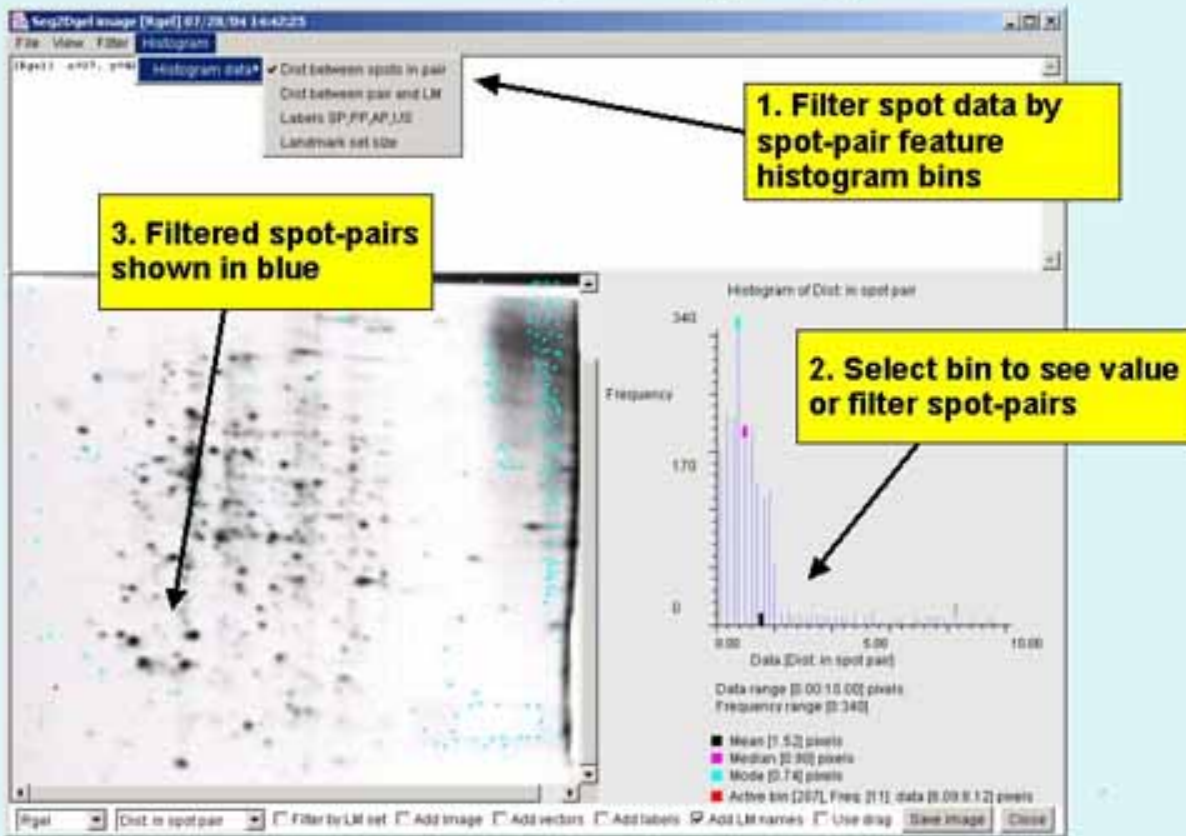
3.11 Image viewer - labels filter LM sets A, B, C, D

Image Viewer - landmark sets "A, B, C, D"



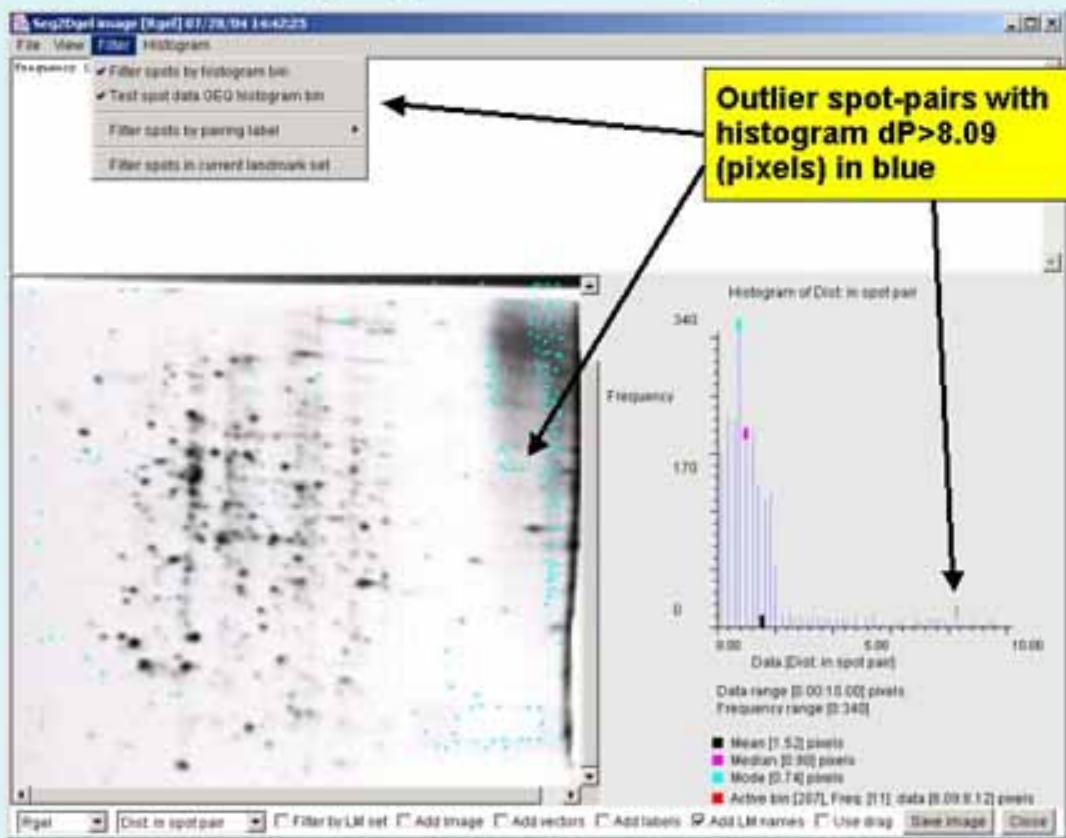
3.12 Image viewer - histogram filter

Image Viewer - histogram of pairing features



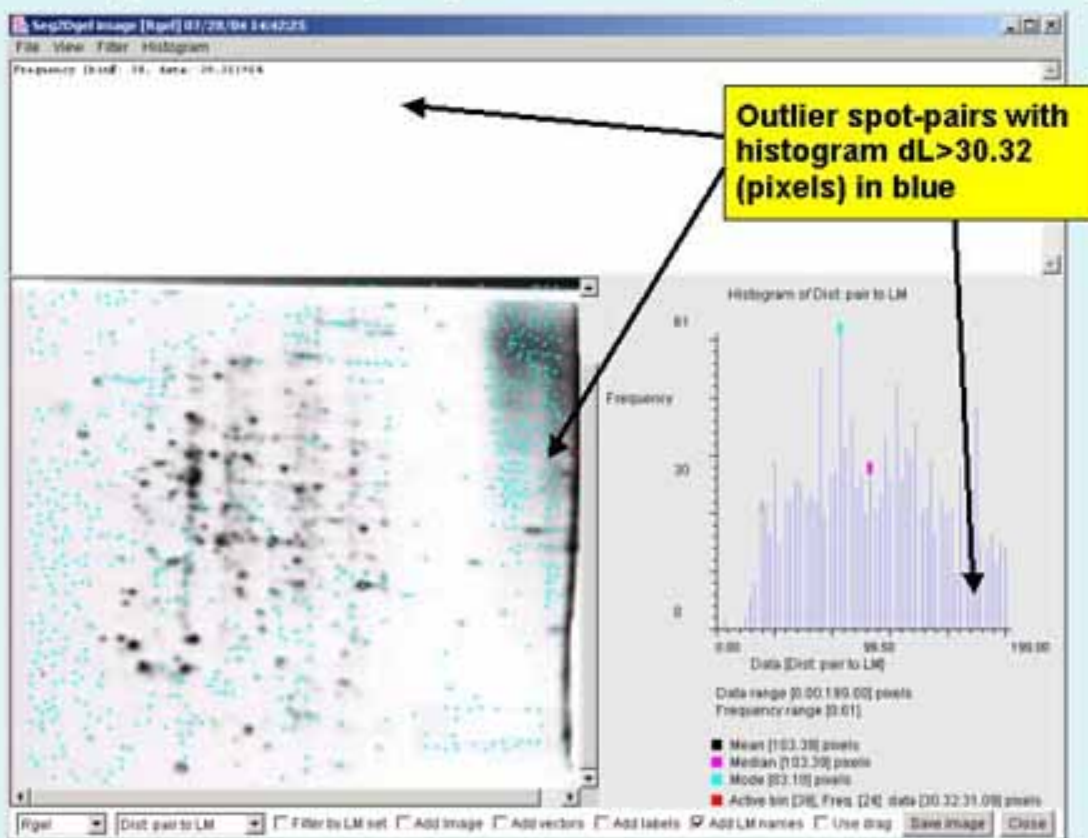
3.13 Image viewer - histogram filter by dP limit

Image Viewer - spot-pairs filtered by dP, dist. Between spots



3.14 Image viewer - histogram filter by dL limit

Image Viewer - spot-pairs filtered by dL, dist. LMs to pairs



[Contact us](#)

CmpSpots is a contributed program available at open2dprot.sourceforge.net/CmpSpots

Revised: 10/04/2005

CmpSpots Jar Version

The CmpSpots program is available on the Open2Dprot server
<http://open2dprot.sourceforge.net/CmpSpots>

or from the mirror server at
<http://www.lecb.ncifcrf.gov/Open2Dprot/CmpSpots>

After you have downloaded and installed the program, you can update it quickly by having CmpSpots copy the CmpSpots.jar file from the Web server using the (File | Update from Web server | CmpSpots program) menu command. After you do the update, you need to restart the program to use the new version.

Until further notice, CmpSpots is to be considered pre-alpha-level code until it is officially released. This means that there may be some functionality not fully implemented, that works incorrectly, or that has changed. New commands and functionality are in the process of being added. Please report problems and suggestions to us.

1. Full download installation CmpSpots version 0.2.9


Revised: 12-21-2005

2. Update CmpSpots program (CmpSpots.jar) version 0.2.9

Revised: 12-21-2005

CmpSpots - (PDF) documents

There are a number of Adobe Acrobat PDF formatted documents and slide shows available for the CmpSpots project.

If you do not have Adobe Acrobat, you can download it for free from Adobe.  The following lists some of the PDF documents you may download. As others are created, they will be added to this list.

1. [CmpSpots Overview slide show](#) for use as a printable document.
2. [CmpSpots Overview slide show](#) for use as a printable document (2 panels/page).
3. [CmpSpots Overview slide show](#) for use as a printable document (6 panels/page).
4. [This CmpSpots Web site as a PDF document](#) for use as a printable document

[Contact us](#)

CmpSpots is a contributed program available at open2dprot.sourceforge.net/CmpSpots

Revised: 110/04/2005 004

CmpSpots Revision History

This describes the revision history of the CmpSpots program of released or soon to be released versions. Generally, only the most recent versions are kept on the Web site (see [Version](#)). Documentation for unresolved bugs and requested functionality is found in the [status](#) Web page.

- **V.0.2.10 05-31-2006:** Get (File | Updates) from main server jar/, demo/, dtd/, scripts/ etc. directories.

- **V.0.2.9 12-21-2005:** Used more robust LMset/DbLM code and read TIFF before GIF or JPEG.
- **V.0.2.6 08-05-2005:** Fixed several bugs in zoom; 6x zoom filtered spots were incorrect; scroller alignment did not work. When Image Viewer was closed it would not continue correctly when restarted.
- **V.0.2.4 05-29-2005:** Fixed bugs in Image Viewer when change magnification and also histogram spot filtering.
- **V.0.2.3 05-29-2005:** Fixed bugs in XML SPF writer in the O2Plib.db.DbPairSamples.
- **V.0.2.2 05-27-2005:** Fixed bugs in Image Viewer GUI.
- **V.0.2.1 05-09-2005:** Fixed bugs in AP and US pairing that occurred a small percentage of the time. Also modified the SPF XML schema in the O2Plib.db.DbPairSamples and DbPspot library. Fixed display of spot pairs for Sample in the popup viewer (it previously only displayed those in the Rsample). Modified the font size for labeling in the image in the popup viewer as the zoom factor increased.
- **V.0.1.16 04-20-2005:** Fixed bugs in library and synchronized with O2Plibrary. Added Zoom feature and changed layout of Image Viewer popup slightly.
- **V.0.1.12 03-21-2005:** Fixed bug in -usage and -version switches. Changed SSF input files so it looks for <sample-name>-SSL.xml rather than <sample-name>.xml.
- **V.0.1.7 01-18-2005:** Cleanup and sync with O2Plib.
- **V.0.1.6 12-20-2004:** Improved secondary pairing.
- **V.0.1.5 12-17-2004:** Fixed pairing errors introduced when refactored the code during cleanup.
- **V.0.1.3 12-08-2004:** Cleanup and refactored code using new O2Plib modules. Fixed bugs which improves pairing. This may result in new pairing results.
- **V.0.1.1 11-22-2004:** Cleanup and refactored code so uses new O2Plib methods. Added new switch -dtd (default is -nodtd) to add (omit) Open2Dprot-SPF.dtd when generate the SSF files.
- **V.0.0.9.1 10-29-2004:** Cleanup the manual and library update.
- **V.0.0.9 10-12-2004:** Now using the Open2Dprot library O2Plib instead of all code being in CmpSpots. Also changed the XML schema and DTD.
- **V.0.0.8 08-25-2004:** enabled the command line processing using the CmpSpots.exe file for running it in batch. Run the command with "-nodemo" and other switches you need as well as the gel image or gel sample name (if using the -accession file option). A new switch -default was added to give you the -demo options but with -nogui and with no -rsample and no -sample options prespecified.

- **V.0.0.7 08-11-2004:** The new SAX XML readers work for SSF, accession and landmark XML input files. The new default is set to inputFormat:X,X,X to read all three files as XML.
- **V.0.0.6 08-10-2004:** Changed accession database input file format so that the sample name is now "Sample" instead of "PubPixFileName". Changed landmark database input file format so that the Sample and Rsample names are now "Sample" and "Rsample".
- **V.0.0.4 08-05-2004:** Generalized the program to other types of 2D data where the objects used in an experiment are a set of samples rather than just a set of gels. Refactored code by renaming all references to (Gel Segmentation File, GSF, gsf) to (Sample Spot-list File, SSF, ssf), and (Gel Comparison File, GCF, gcf) to (Sample Paired-spot list File, SPF, spf) including both the Java code, DTD and documentation. NOTE: This must be synced with changes in the Seg2Dgel subproject. Fixed misc. bugs with command line options GUI editor, refresh with Image Viewer, re-initialization problems when rerunning it multiple times or after changing the command line options.
- **V.0.0.3 07-30-2004:** Generalized the program to other types of 2D data. Refactored code by renaming all references to (Rgel and gel) to (Rsample and Sample).
- **V.0.0.2 07-29-2004:** Cleanup and fixed some of the XML readers with Epilogue errors. Added -inputFormat:{gsfMode,accMode,lmsMode}.
- **V.0.0.1 07-28-2004:** This is a pre-alpha version. The data is hardwired for now. The histogram is not fully working. It requires tab-delimited accession and landmark database files for now. It will accept either -gsfFormat:X or -gsfFormat:F GSF files. It will generate either -gcfFormat:X or -gcfFormat:F GCF.y

y

[Contact](#)
[us](#)

CmpSpots is a contributed program available at open2dprot.sourceforge.net/CmpSpots

Revised: 05/31/2006

CmpSpots Program Status

The following are known bugs, suggested features, functions that are being debugged or being developed. Documentation for resolved bugs is found in the [revision history](#) Web page.

- [Revision history](#)
- [FAQ](#)
- [Known bugs](#)
- [Suggested Features/FAQ](#)

Known Bugs

The neumonic *GGGR* is shorthand for "Grep, Get Gestault and Refactor".

- **V.0.0.4 08-05-2004:** Histogram data selections are not working for the "Lbls: SP,PP,AP,US" and for "Sizes: LM sets".
- **V.0.0.4 08-05-2004:** The "CmpSpots.properties" file in the current project directory does not get saved or read properly at this time.

Suggested Features

- **V.0.0.1 05-03-2004:** [TODO]

[Contact us](#)

CmpSpots is a contributed program available at open2dprot.sourceforge.net/CmpSpot

Revised: 10/04/2005

CmpSpots Program Frequently asked questions (FAQ)

The following are some frequently asked questions. This is a list of some of the questions that we will be answering. Please [send us your questions](#) about CmpSpots so we can add them and the answers to this list.

- [Revision history](#)
- [Known bugs](#)
- [Suggested Features/FAQ](#)

FAQ

- How do I specify the two samples to pair? Through the **-rsample:** and **-sample:** command options.

- What is a project and a project directory? What goes into this directory and can I have more than one? The project is where all of the input and output files are stored in your computer.
- How do I change the options? If you are running the -gui option (Graphical User Interface or GUI). Invoke the "Edit options" to edit the new parameters, then press "Pair spots" to re-pair the data with the new parameters.
- What is a command line switch? How is it used? You specify the options to CmpSpots through command line switches. If you are running the GUI version, then you specify the command line switches through the Options Wizard window by pressing the **Edit options** button in the Report window.
- How do I adjust the spot pairing parameters? Using the -gui option. Invoke the "Edit options" wizard to edit the new parameters, then press "Pair spots" to re-pair the data with the new parameters.
- Can I save the options that I have just changed and apply them to other samples pairing operations? They are saved in the startup file "CmpSpots.properties" when you exit the program and are read when you start the program again.
- How do I handle requiring different sizing parameters for different samples? Using the -gui option. Invoke the "Edit options" to edit the new parameters, then press "Pair spots" to re-pair the data with the new parameters.
- How do I see the results of a spot pairing? When spot pairing is finished, invoke the "Image Viewer".
- How do I stop the spot pairing program when it seems to have gotten stuck or is taking too long? Press the "Stop pairing" button.
- How do I generate an XML output file? The default is to use the "-spfFormat:X" switch.

[Contact
us](#)

CmpSpots is a contributed program available at [open2dprot.
sourceforge.net/CmpSpots](http://open2dprot.sourceforge.net/CmpSpots)

Revised: 10/04/2005