DNAinsight: An Image Processing System for 2-D Gel Electrophoresis of Genomic DNA

Katsutoshi Takahashi [1]
sltaka@infor.kanazawa-it.ac.jp

Masayuki Nakazawa [2]
nakazawa@nakazawa.ishikawa-pc.ac.jp

Yasuo Watanabe [1]
watanabe@infor.kanazawa-it.ac.jp

[1] Kanazawa Institute of Technology
Ohgigaoka 7-1, Nonoichi, Ishikawa 921, Japan

[2] Ishikawa Polytechnic College
Yuigaoka I 45-1, Anamizu, Ishikawa 927, Japan

Abstract

We have developed a powerful image processing system DNAinsight, which performs automated detection of several thousands of spots found on autoradiogram images obtained with 2-D gel electrophoresis of genomic DNA. Algorithms and parameters for detecting spot locations and intensities are carefully chosen so as to enable reliable and rapid processing of 2-D gel electrophoretograms based on the RLGS (restriction landmark genomic scanning) method. In DNAinsight, matching of several related spot patterns, such as those from tumor-cell and normal-cell, can be accomplished rapidly with easy operations, being solved by comparing the Delaunay net and relative neighborhood graph. The automated and accurate image processing system strongly supports the rapid identification and analysis of genetic variation in the DNA of humans and other animals.

1 Introduction

In the passed a few decades, researches on genetic analysis have revealed the relation between genetic alteration and several diseases, such as tumor and hereditary diseases. Identification of pathogenic genes contributes not only to progress on DNA diagnosis, but also to elucidation of complex gene regulation mechanisms.

Positional cloning is one of the most powerful experimental strategy to clone pathogenic genes [1, 4]. At the beginning of this strategy, it is necessary to determine and position the genetic locus of the pathogenic gene, based on linkage analysis and/or observed chromosomal aberrations. Molecular biological genome scanning is useful for determining the candidate genetic loci on the complex genomic DNA. In principle, the genome scanning is defined as overall detection of the physical condition of whole genomic DNA. To scan whole complex genomic DNA, however, only landmark information is detected, ignoring the other genomic DNA regions. One key step in high-speed identification and genetic analysis of pathogenic genes is the high-speed and highly reproductive technique for survey of presence or absence of landmarks throughout a genome, and for measurement of their copy number in each locus.

Southern hybridization-based method [3, 13] and PCR-based method [11] have been developed and applied to scan genomic DNA. In Southern hybridization using a unique probe and in PCR for amplifying a unique sequence, only one locus can be assayed in one procedure. In hybridization-based and PCR-based multiplex methods, such as DNA fingerprinting using repeating sequences as probes, multiple loci can be screened in one procedure. However, the scanning field and scanning speed are limited by the type of available repeating sequences.

Recently, Hatada et al. introduced a new concept , termed "restriction landmark," in which each restriction enzyme recognition site can be used as a landmark [7]. Based on the concept, RLGS (Restriction Landmark Genomic Scanning) method was developed [8]. This method employs direct end-labeling of the genomic DNA digested with a restriction enzyme and high-resolutional two-dimensional electrophoresis. It provides an alternative multiplex approach to genome analysis that can be used to map a large number of loci simultaneously. RLGS method can yield several thousands of landmark spots on a two-dimensional electrophoretogram in one procedure, and the landmark spot intensity reflects the copy number of the restriction fragment. Thus, RLGS method has a great advantage in capacity and speed for scanning of an entire genome compared with the conventional hybridization-based or PCR-based methods and has been applied for various genetic analysis [9, 10].

Applying RLGS method to several related genomic DNA and comparing resultant 2-D electrophoretograms, one can detect molecular changes, such as deletions, additions, amplifications or DNA methylations, that occur at or near to the restriction enzyme recognition sites. It is difficult and sometimes impractical, however, either to separate all of the several thousands of spots or to determine their intensities by visual inspection. It is also difficult by hand to compare several distorted complicated images. Such difficulties in the image analysis make the high-speed multiplex genome scanning impractical. To overcome the above problems, we have developed a powerful image processing system dealing with two-dimensional electrophoretograms of genomic DNA, called **DNAinsight**. It offers automated extraction of several thousands of spots and the ability to detect differences in spot locations and intensities among several related images, allowing one to perceive the translocations, amplifications and deletions of landmark spots.

2 Basic problems on image analysis of RLGS profiles

RLGS method begins with cleavage of purified genomic DNA by a rare restriction enzyme (A) such as *Not*I. The cleavage sites are end-labeled and the labeled DNA fragments are further digested by restriction enzyme (B) such as *Eco*RV, which recognition site appears more frequently in the genome than of restriction enzyme A. The purpose of this second cleavage step is just the reduction of the DNA fragment in its length, and this step is sometimes skipped, depending on the restriction enzyme A used. The mixture of the DNA fragments is then fractioned through a first-dimensional agarose gel electrophoresis. The fractionated DNA fragments are then digested with a restriction enzyme (C) such as *Hinf*I in the gel, and subjected to a second-dimensional polyacrylamide gel electrophoresis. After autoradiography of the dried up gel, DNA fragments having labeled *Not*I termini appears as several thousands of spots on the RLGS electrophoretogram (hereafter referred to as RLGS profile).

Figure 1 shows an example RLGS profile of whole genomic DNA derived from tumor-tissue, giving approximately 1200 landmark spots. This RLGS profile was obtained by employing *Not*I, *Eco*RV and *Hinf*I as restriction enzyme A, B and C, respectively. In this case, 17 inch $\times$ 14 inch sized polyacrylamide gel was required to resolve all of such spots. One spot on the RLGS profile corresponds to one locus and its intensity , defined as integrated optical density, reflects copy number of the locus. In other word, each locus is characterized by spot location (x, y) and its intensity I, letting horizontal location be x and vertical location be y. The coordinates (x, y) correspond to the distance on genomic DNA from the site of restriction enzyme A to B and from the site of enzyme A to C, respectively.

In RLGS profiles, some irregular features in spot shape can be observed frequently. In the above example RLGS profile, the spots located on near right side have long-tailed shape, and some spots show flat shape as shown in figure 2. Besides, a strongly drifted un-uniform background pattern can be seen in the RLGS profile, which may be caused by non-specifically damaged sites on genomic DNA. It is noteworthy that such irregular shapen spots and drifted background seem to be observed more frequently in RLGS profiles of genomic DNA derived from human-tissue, compared to those from cultured cells. Because of such irregularities in spot shape, the conventional spot detection algorithm [6, 14] can not be applied to analysis of RLGS profiles. The development of powerful spot detection

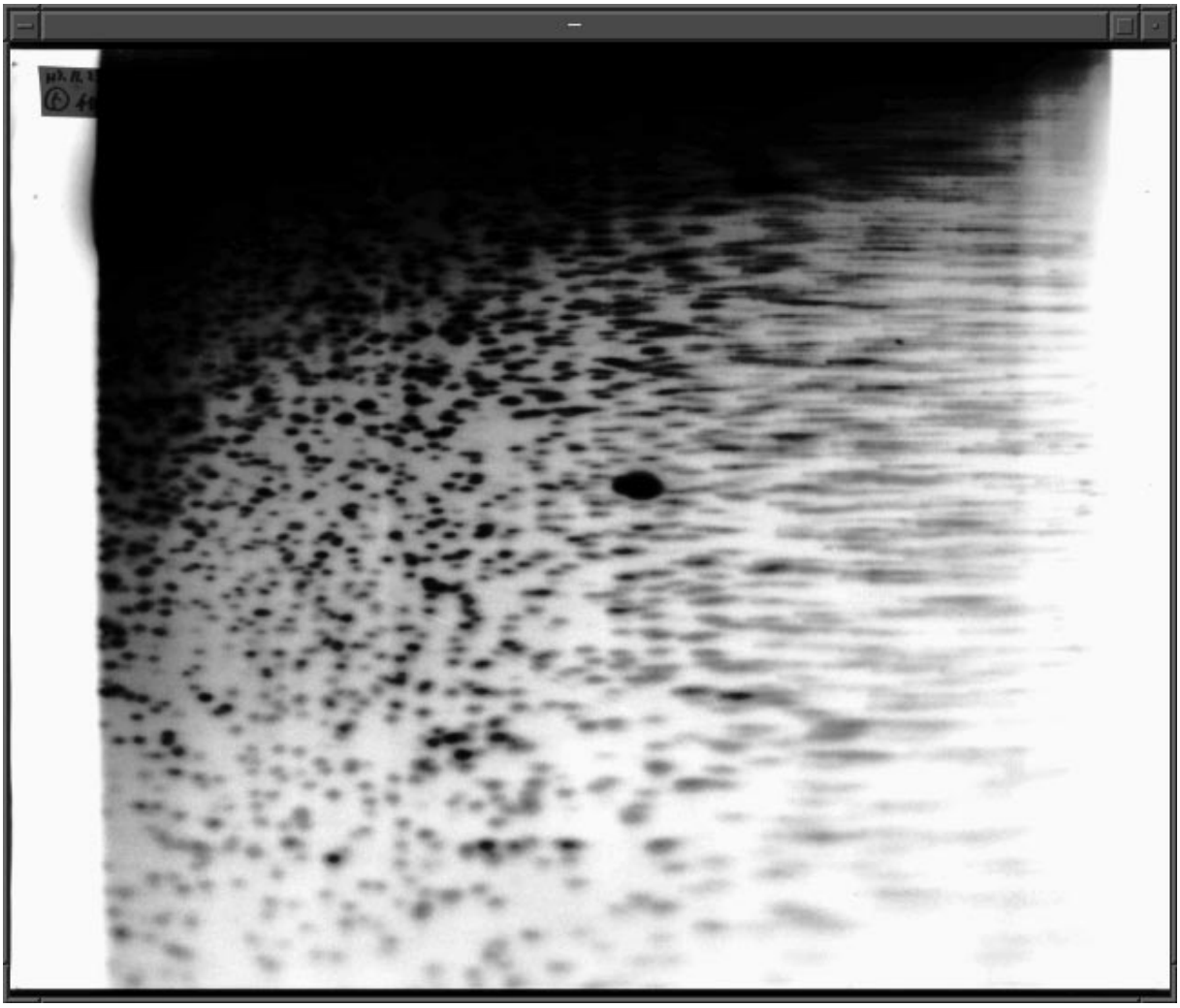


Figure 1: An example RLGS profile of whole genomic DNA derived from tumor-tissue. The upper black-outed region corresponds to the *Not*I site end-labeled DNA fragments which were not digested in the third enzyme cleavage step.

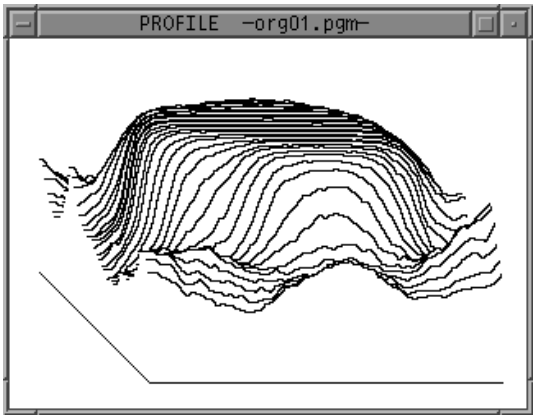
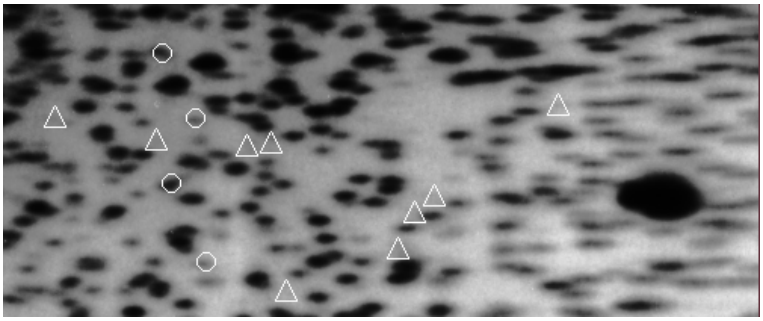
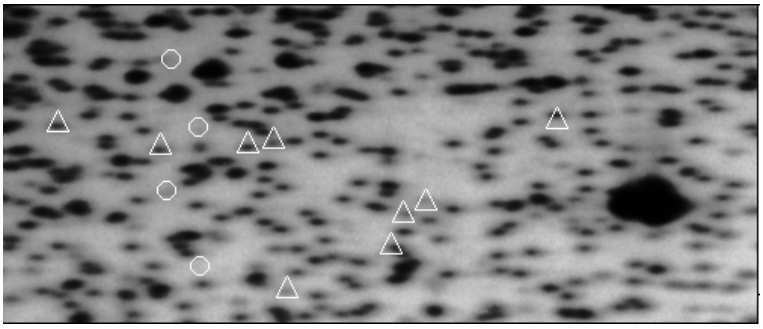


Figure 2: 3D-profile of a landmark spot, where the flat shape is caused by signal saturation.



(a)



(b)

Figure 3: The difference in two RLGS profiles, one derived from human tumor-tissue (a) and another from human normal-tissue (b).

algorithm, which allows such strong irregularities, is highly desirable. Therefore, the first problem on image processing of RLGS profile is to detect location of a spot and its intensity (x, y, I) accurately, even for a irregular shapen spot.

As described above, molecular changes occurring at or near to the restriction sites can be detected by comparing several related RLGS profiles. In figure 3, molecular changes of genomic DNA occurring in tumor-cell can be observed as differences in the RLGS profiles. In this figure, the symbol $\bigcirc$ indicates a spot which appears in the RLGS profile (a) but disappears in the RLGS profile (b) , though the symbol $\triangle$ means a spot which disappears in (a) but appears in (b). As can be seen from this figure, the RLGS profiles take non-linear distortions and some variations of the spot size and intensity. In general, the whole patterns of the RLGS profiles do not coincide, even if the profiles are derived from exactly the same genomic DNA fragments. The RLGS profiles of humans and other animals are so complicated and disordered that it is difficult to match the corresponding spots on the related RLGS profiles either by hand or by applying the conventional image matching algorithms [2, 5, 12, 14].

Some of the landmark spots might appear on all of the related RLGS profiles. When such spots were known in advance, the knowledge helps to match their corresponding spots on a newly determined RLGS profile. However, such knowledge about a series of the related RLGS profiles is of no use when a profile is derived from genomic DNA of different species or solved by the different restriction enzymes. Thus, the second problem on analysis of RLGS profiles is to recognize the differences in the disordered RLGS profiles, without any prior knowledge of the spot patterns commonly found on a series of related RLGS profiles.

3 Automated recognition of spot locations and intensities

Prior to image processing, X-ray films should be digitized with film scanner. In the above example RLGS profile, 17 inch × 14 inch sized X-ray film was required to resolve over thousand landmark spots, and gives the image file of 5100 × 4200 pixels with scanning at a resolution of 300 dot per inch. The image file was then re-sampled as to give 1275 × 1050 pixels with 8-bit per pixel. The re-sampled image file was, then, provided for proceeding image analysis.

The automated recognition of spot locations consists of the following steps:

(1) Preprocessing for image enhancement and smoothing. Let the preprocessed image be $\phi(x, y)$.

(2) Applying background normalization operation [14] to $\phi(x, y)$. Let the resultant image be $\psi(x, y)$. Then the overall background level t of $\psi(x, y)$ was determined by means of conventional smoothed density histogram method. The binary image $f(x, y)$ of $\psi(x, y)$, which gives the possible spot domains, was obtained as

$$f(x,y) = \begin{cases} 1, & \text{if } \psi(x,y) \geq t \\ 0, & \text{otherwise} \end{cases} .$$

(3) Applying a ring operator to $\phi(x, y)$ on the domain $\{(x, y)|f(x, y) = 1\}$ to detect the local maxima independently of background density. Suppose that

$$C(x,y) = \{(u,v)|(u-x)^2 + (v-y)^2/\alpha^2 \leq r_M^2\}$$

and

$$R(x,y) = \{(u,v)|r_m^2 \leq (u-x)^2 + (v-y)^2/\alpha^2 \leq r_M^2\}$$

where α is the ratio of the minor axis to major axis for an ellipse. Then the output of a ring operator is defined by

$$h(x,y) = \max_C \phi(x,y) - \max_R \phi(x,y).$$

Here, the elliptic ring operator was adopted in order to allow efficient detection of flat shapen spots which have long tails in the first dimensional electrophoresis axis x.

(4) Labeling the candidate spots detected in the above step for their identification.

Figure 4 shows the spots detected by the above steps on the RLGS profile of 660 × 665 pixels. In this example, almost all of 1065 landmark spots were recognized correctly, with only 48 spots not detected and 33 wrongly recognized spots. One of the most important things in spot detection is what kind of operator should be selected as a spot detector. As shown in table 1, the ring operator has much power in recognizing landmark spots on RLGS profile, compared to two types of other operators, i.e., the usual simple peak detector combining horizontal and vertical scans and the matched filter of Gaussian type.

Though the ring operator shows better performance than the other, it still has several tens of spots wrongly recognized and not detected. It is, however, relatively easy to rectify such amount of false positives and negatives. We provide several utilities which help one to perceive spots ill-recognized or should be recognized, together with the tools for revising such spots manually or semi-automatically, as described in later section.

Once a spot location was recognized, its intensity should be calculated in order to estimate copy number of the landmark. Occasionally, a pair of landmark spots overlaps each other in RLGS profile such that the spots cannot be distinguished clearly. To calculate integrated optical density of such overlapping spots separately, gray levels of the preprocessed image $\phi(x, y)$ are carefully analyzed and each pixel in the image is classified into one of the spot domains with the following steps:

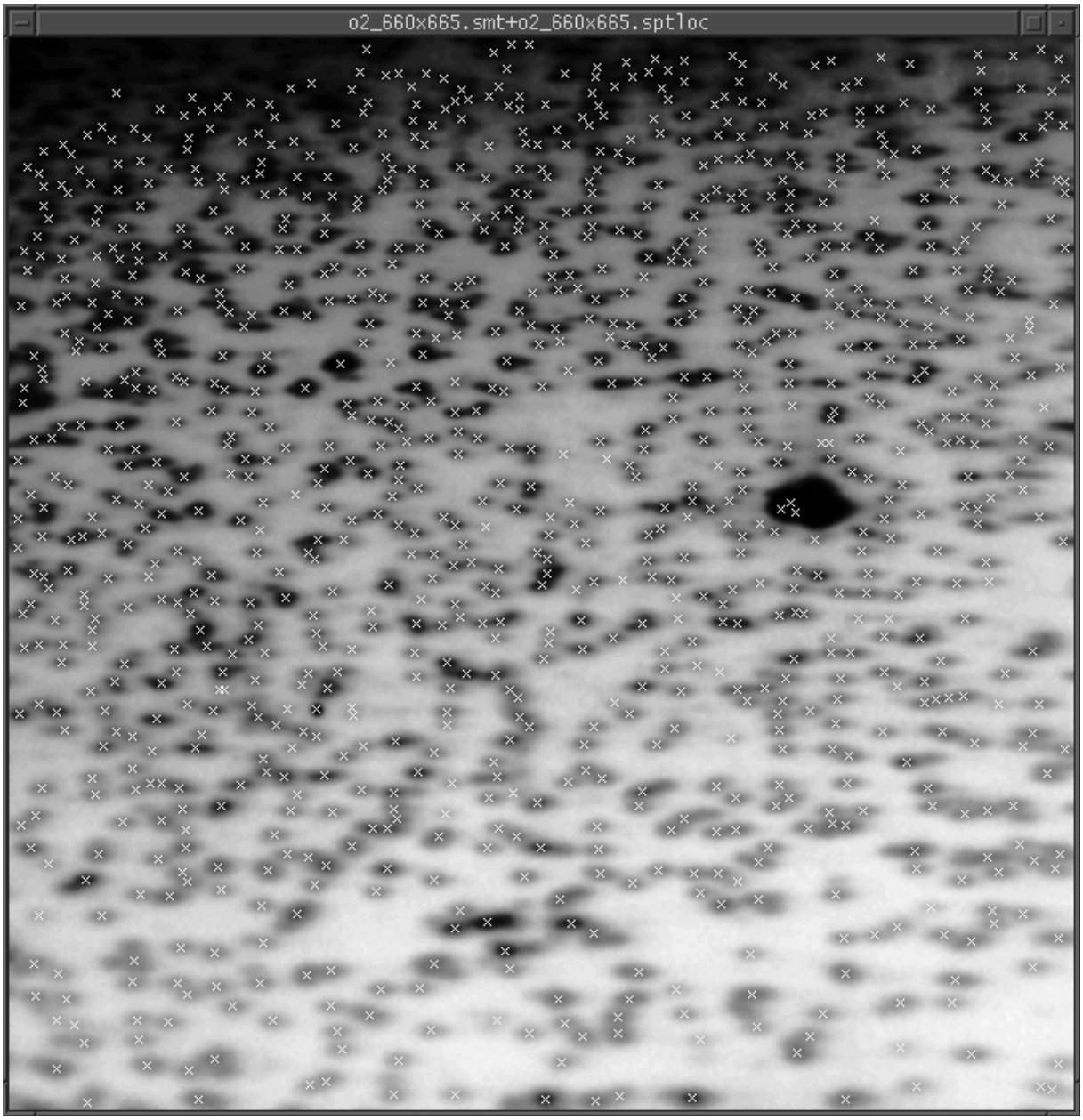


Figure 4: Automatically detected landmark spots on RLGS profile.

Table 1: Number of spots detected by various spot detectors.

Spot detector	Automatically detected spots	Wrongly detected spots	Spots not detected
Peak detector	1281	365	149
Gaussian 1 [1]	899	34	200
Gaussian 2 [2]	836	31	260
Ring operator [3]	1052	33	48

[1] Gaussian matched filter with $\sigma = 2$.

[2] Gaussian matched filter with $\sigma = 5$.

[3] Ring operator with $r_m = 3.5$, $r_M = 5.5$ and $\alpha = 0.8$.

(1) Labeling pixels in the preprocessed image $\phi(x, y)$ whose gray levels are the highest through the image. The labeled pixels are either local maxima or in the regions of local flatness inside the flat shapen spots. Each labeled pixel is then classified into the spot domain D_i, which consists of adjoining pixels. Here, pixels on the domain $\{(x, y)|\psi(x, y) \leq t\}$ are ignored.

(2) A second label marks unlabeled pixels whose gray levels are the highest among all of the unlabeled pixels. The second labeled pixels are then classified into their adjoining spot domains which appeared in the previous steps, with enlarging each spot domain. This classification is continued until no more second labeled pixels adjoin any spot domains. Then, the residual unclassified pixels are offered to define new spot domains.

(3) The above step is repeated until any unlabeled pixels cannot be found on the image.

When a spot domain D_i computed in the above steps is assigned to a spot, the intensity I of the spot is defined as

$$I = \sum_{(x,y)\in D_i} \psi(x, y).$$

4 Structured features and matching of spot patterns

The next problem in analysis of RLGS profile is matching two or more RLGS profiles in order to detect such differences as translocation, amplification and deletion of landmark spots. RLGS profiles, however, cannot be compared by a simple image overlay technique, because they take non-linear distortions and some variations in spot size and intensity. In our application, instead, it is necessary to consider a matching problem of featured point patterns represented by each RLGS spot location and intensity (hereafter referred to as RLGS patterns). Various methods of two-dimensional pattern matching have been proposed so far [2, 12, 14], but these are time-consuming and highly depend on the initial state of iteration. Our approach to the point pattern matching deals with correspondence of two structured graphs, that is, Delaunay net and relative neighborhood graph (hereafter referred to as DN and RNG, respectively,) rather than two sets of points.

It is known that RNG is a subset of DN and also a superset of minimal spanning tree [15], and that RNG is insensitive to geometric distortion while DN is sensitive. Therefore, we adopt an algorithm in which the RNG of the reference RLGS pattern is used as a guide for matching any nodes in the DN of object RLGS pattern against those in the reference RNG, though the reference RNG is not an exact subgraph of the object DN.

With initial points in both graphs given, the graph matching starts with these initial points and progresses with breadth-first search, minimizing the following evaluation function for the depth k.

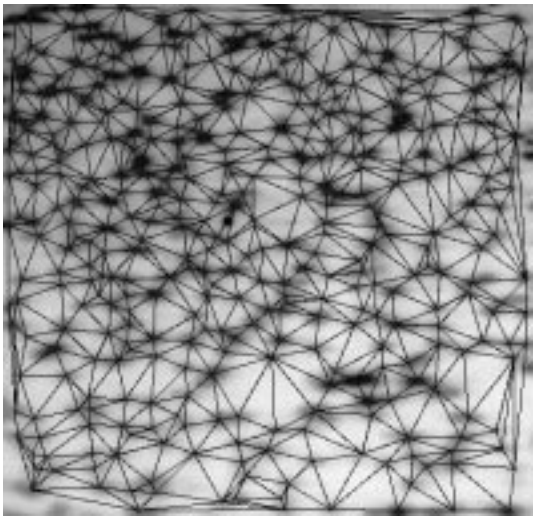
$$E_k = \sum_i \Big[\{1 - S(\mathbf{a}_i^{(k)}, F(\mathbf{a}_i^{(k)}))\}^2 + \{1 - S(P_r(\mathbf{a}_i^{(k)}), P_o(F(\mathbf{a}_i^{(k)})))\}^2\Big],$$

where for the vectors $\mathbf{u}$ and $\mathbf{v}$,

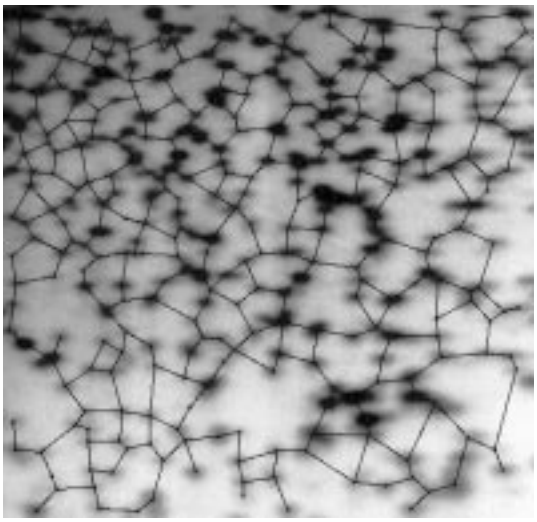
$$S(\mathbf{u}, \mathbf{v}) = \frac{(\mathbf{u}, \mathbf{v})}{|\mathbf{u}||\mathbf{v}|},$$

$\mathbf{a}_i^{(k)}$ is the i-th vector corresponding to the directed arc to be traversed, and F is the mapping from arcs of the reference RNG to those of the object DN. In addition, $P_r(\mathbf{x})$ is the vector representing a referenced subimage centering at the destination point of $\mathbf{x}$ and $P_o(\mathbf{x})$ means the similar vector with respect to the object subimage. Here, the mapping F which minimizes the above evaluation function E_k gives the optimal spot correspondence in the reference and object RLGS patterns.

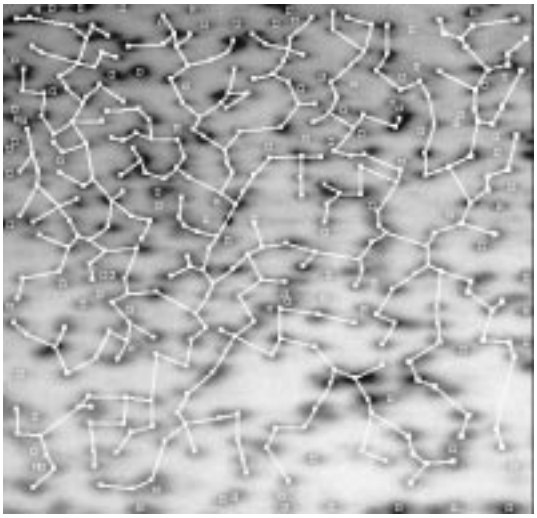
In figure 5, DN of the object RLGS pattern (a) was matched against RNG of the reference RLGS pattern (b). The filled square shown in (c) and (d) indicates a spot which has correspondence on the



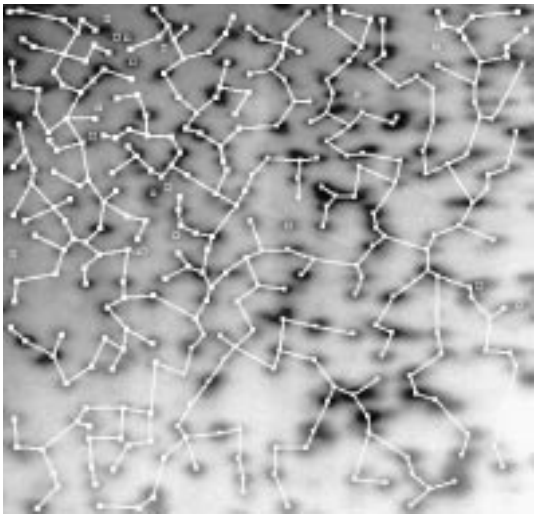
(a) DN (Delaunay net) of the RLGS pattern to be matched.



(b) RNG (relative neighborhood graph) of the RLGS pattern to be referenced.



(c) Matched DN of the object RLGS pattern.



(d) Matched RNG of the reference RLGS pattern.

Figure 5: The object DN and the reference RNG of matched RLGS patterns, being overlayed on the reference and object RLGS profiles, respectively. The filled square in (c) and (d) indicates that a pair of spots corresponds each other in the reference RNG and the object DN, though the open square means that the spot has no correspondence on the other RLGS pattern or no visit during the matching process. Also in (c) and (d), only the matched arcs are drawn with thick solid line.

other RLGS profile, though the open square means that the spot has no correspondence. In addition, the thick solid line shown in (c) and (d) indicates that the arc is matched with the arc in another graph. The reference and object graphs drawn only by the matched arcs are quite useful to confirm the matching result.

In the above example, the reference RNG (d) consists of 334 nodes and 445 arcs, and the object DN contains 412 nodes and 1204 arcs. The computation time, which is the sum of user-time and system-time, required to make the reference RNG (d) and the object DN (c) were only 0.8 sec and 0.9 sec, respectively, on DOS/V computer (Pentium/75MHz with 16MB EDO-DRAM) running with Linux. It is noteworthy that matching of the object DN (a) against the reference RNG (b) took only 0.8 sec on the same computer.

Furthermore, our algorithm can be applied to larger and more disordered RLGS patterns. When the object RLGS pattern, derived from the RLGS profile of 1245 $\times$ 832 pixels with 2104 spots, was matched against the reference RLGS pattern, obtained from the RLGS profile of 729 $\times$ 867 pixels with 1002 spots, 730 spots in each pattern were matched automatically in 3.9 sec (image data not shown). Only 10 spots were mismatched out of the 730 matched spots, while 46 unmatched spots in each RLGS pattern have their corresponding spots on the other.

As demonstrated above, our matching algorithm accurately and rapidly recognizes a pair of the corresponding spots, even if the RLGS profiles are very large and disordered. The automated matching algorithm provides simultaneously the spots which have no correspondence on the other RLGS profile, which are the clue to the elucidation of genetic alterations occurring on the genomic DNA.

5 DNAinsight system

The spot recognition module and the spot pattern matching module are integrated into the powerful image processing system, **DNAinsight**, together with several utility modules as shown in figure 6. Each component module in **DNAinsight** is controlled from main module, so-called **navigator**, which has a functional graphical user interface such that almost of all image analysis operations can be performed by simple mouse operation. The analysis flow of RLGS profiles in this system is as follows:

(*i*) Starting the **navigator** with un-cropped RLGS profile image. The image is then displayed in the main window.

(*ii*) Setting the object regions on the displayed image by simple mouse dragging. The cropped subimage corresponding to an object region is then provided for further analysis.

(*iii*) Performing the first stage of spot recognition process by choosing an item from the pull-down menu. In this stage, preprocessing, background normalization and binarization of the cropped subimage are carried out automatically, followed by the recognition of spot locations by means of the ring operator. At the end of this stage, the automatically recognized spot locations are stored in a text file.

(*iv*) Confirming and revising the automatically recognized spot locations, if necessary. Such functional module, so-called **revisor**, can be activated from a pull-down menu item of **navigator**. The **revisor** displays the recognized spots, overlaying them onto the cropped subimage, with or without various image enhancements. Then, one can delete or mark spots manually on **revisor** window, while the spot location, not recognized automatically in previous step, can be detected semi-automatically. Besides, support utility modules, such as **contour-viewer**, **3D-profiler** and **spotdomain-viewer**, help to analyze spot shape in detail, perceiving spots ill-recognized or should be recognized.

(*v*) Computing intensity of each spot.

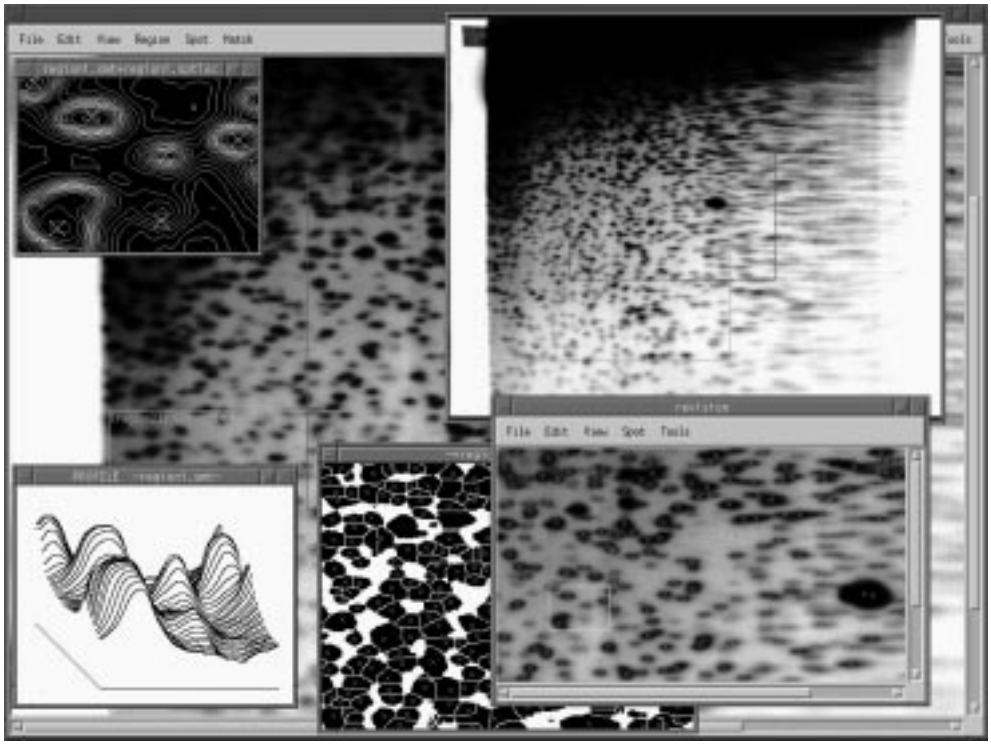


Figure 6: **DNAinsight** system.

(*vi*) Matching the RLGS pattern against the reference RLGS pattern. Such functional module, so-called **matchmaker**, can also be activated from a pull-down menu item of **navigator**. Once it is activated it displays both of the reference and object RLGS profiles, then urges to select a pair of initial spots from which matchmaking starts. At the end of this step, **matchmaker** stores the resultant spot matching table into a text file. In the spot matching table, spots on the two profiles are classified into three categories, i.e., "matched spots," "appeared spots" which present only on the object profile and "disappeared spots" which disappeared from the reference profile.

Through the above steps, one can perceive differences in two RLGS profiles without any prior knowledge of a series of the related RLGS patterns, which is necessary to analyze RLGS profiles by visual inspection. To elucidate the linkage between the landmark spots and their genetic information, further systematic analysis should be performed on a series of the spot matching tables each of which is obtained by matching the related RLGS pattern against a reference pattern.

DNAinsight was implemented on DOS/V computer compatible to PC running with Linux. Linux is a OS (Operating System) almost compatible to UNIX and is widely available without any fee. It is distributed as CD-ROM media, and can be obtained even via computer network.

Each component module was carefully designed and implemented as to allow its operation on computer either with slow CPU or less memory capacity. The component modules are efficient enough to process large sized RLGS profiles even on cheaper DOS/V computer, as shown in table 2. In principle, they can be implemented on any computers running with UNIX, on which X-window and Motif are available, from personal computers to engineering workstations.

Table 2: Computation time [1] (in sec) required to process digitized RLGS profile.

Image size	preprocessing	background [2]	binarization	ring [3]	ring [4]	integration	total
271 × 229	0.2	0.3	0.4	1.0	1.7	5.6	9.2
361 × 284	0.2	0.5	0.5	1.9	3.2	7.9	14.2
614 × 468	0.5	1.2	1.4	4.1	7.5	23.1	37.8
660 × 665	1.0	1.8	2.1	11.1	14.8	40.9	71.7

[1] User-time + system-time, which was measured on DOS/V computer compatible to PC running with Linux: Pentium/75MHz with 16MB EDO-DRAM.

[2] Background normalization.

[3] Ring operator with $r_m = 3.5$, $r_M = 5.5$ and $\alpha = 0.8$.

[4] Ring operator with $r_m = 5.5$, $r_M = 7.5$ and $\alpha = 0.9$.

6 Summary

We have developed a powerful image processing system, **DNAinsight**, which deals with autoradiogram images obtained with two-dimensional electrophoresis of genomic DNA. **DNAinsight** provides a rapid and objective way to analyze several thousands of landmark spots found on large sized RLGS profiles and also to extract matched, appeared or disappeared spots from the related RLGS profiles. Actually, it takes about ten minutes to detect all spots on a RLGS profile and to compare them with those on a reference RLGS profile even with cheaper DOS/V computer running with Linux, while it takes about two or three hours by visual inspection.

As a conclusion, **DNAinsight** highly contributes the automation of the high-speed multiplex genome scanning. It should realize the truly high-speed and highly reproductive genetic analysis, which elucidate the linkage between landmark loci and their genetic information, in concert with the progress on the experimental techniques.

7 Acknowledgments

We are grateful to Prof. Ueda at Keio University for providing several RLGS autoradiogram images, and to Dr. Hirota at NBT Inc. for his technical supports. We would also like to thank Prof. Sugihara at Univ. of Tokyo for his kindly providing us a computer program which generates Voronoi Diagram. His program was used to verify our Delaunay net production algorithm. Finally, we acknowledge Dr. Hayashizaki at The Institute of Physical and Chemical Research (RIKEN) for his kind suggestions on the first stage of this work.

References

[1] Ballabio, A., "The rise and fall of positional cloning?," *Nature Genet.*, 3:277-279, 1993.

[2] Barnard, S.T., Thompson, W.B., "Disparity Analysis of Images," *IEEE Trans. Pattern Analysis and Machine Intelligence, PAMI-2*, 4:333-340, 1980.

[3] Brilliant, M.H., Gondo, Y., Eicher, E.M., "Direct molecular identification of the mouse pink-eyed unstable mutation by genome scanning," *Science*, 242:566-569, 1991.

[4] Collins, F., "Positional cloning: let's not call it reverse anymore," *Nature Genet.*, 1:3-6, 1992.

[5] Fischler, M.A., Elschlager, R.A., "The representation and matching of pictorial structures," *IEEE Trans. Computer, C-22*, 1:67-92, 1973.

[6] Garrals, J.I., "Two-dimensional Gel Electrophoresis and Computer Analysis of Proteins Synthesized by Clonal Cell Lines," *J. Biological Chemistry*, 254:7961-7977, 1979.

[7] Hatada, I., Hayashizaki, Y., Hirotsune, S., Komatsubara, H., Mukai, T., "A genomic scanning method for higher organisms using restriction sites as landmarks," *Proc. Natl. Acad. Sci. U.S.A.*, 88:9523-9527, 1991.

[8] Hayashizaki, Y., Hirotsune, S., Okazaki, Y., Hatada, I., Shibata, H., Kawai, J., Hirose, K., Watanabe, S., Fushiki, S., Wada, S., et al., "Restriction landmark genomic scanning method and its various applications," *Electrophoresis*, 14:251-258, 1993.

[9] Hayashizaki, Y., Hirotsune, S., Okazaki, Y., Shibata, H., Akasako., A., Muramatsu, M., Kawai, J.,Hirasawa, T., Watanabe, S., Shiroishi, T., et al., "A genetic linkage map of the mouse using Restriction Landmark Genomic Scanning (RLGS)," *Genetics*, 138:1207-1238, 1994.

[10] Hirotsune, S., Hatada, I., Komatsubara, H., Nagai, H., Kuma, K., Kobayakawa, K., Kawara, T., Nakagawara, A, Fujii, K., Mukai, T. and Hayashizaki, Y., "The new approach for detection of amplification in cancer DNA using Restriction Landmark Genomic Scanning method," *Cancer Res.*, 52:3642-3647, 1992.

[11] Nelson, D.L., Ledbetter, S.A., Corbo, L., Victoria, M.F., Ramirez-Solis, R., Webster, T.D., Ledbetter, D.H., Caskey, C.T., "Alu polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA source," Proc. Natl. Acad. Sci. U.S.A., 86:6686-6690, 1989.

[12] Ranade, S., and Rosenfeld, A., "Point Pattern Matching by Relaxation," *Pattern Recognition*, 12:269-275, 1980.

[13] Southern, E.M., "Detection of Specific Sequences among DNA Fragments Separated by Gel Electrophoresis," *J. Mol. Biol.*, 98:503-517, 1975.

[14] Sternberg, S.R., "Biomedical Image Processing," *IEEE Computer*, January:22-34, 1983.

[15] Toussaint, G.T., "The relative neighborhood graph of finite planar set," *Pattern Recognition*, 12:261-268, 1980.