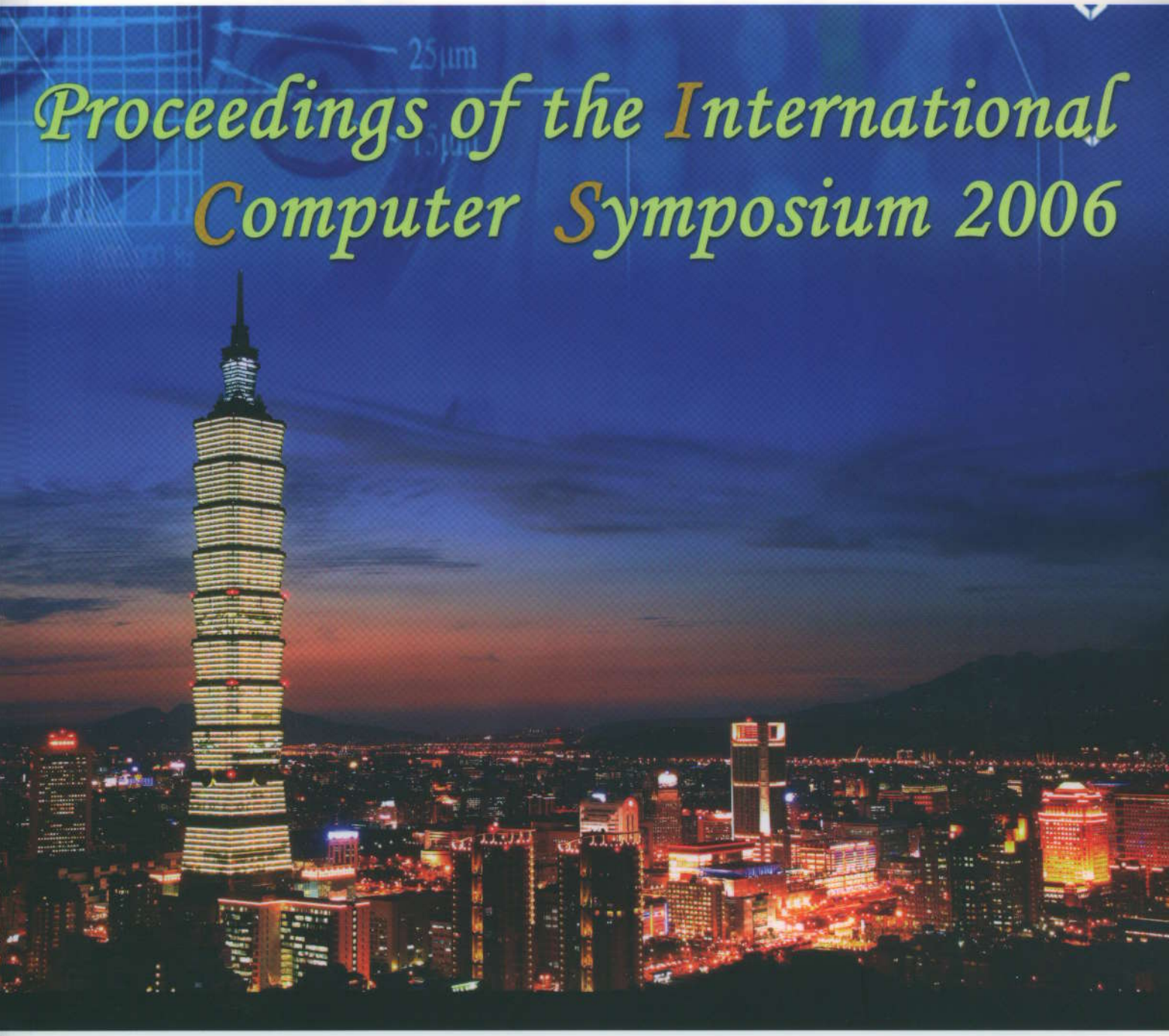
ICS 2006

III



Proceedings of the International Computer Symposium 2006

Taipei, Taiwan, Dec 04-06. 2006

Genotyping Polymorphic Bands of Microsatellite DNA with Automatic Object Segmentation and Translated Absolute Size

Chun-Fan Chang[1,8,*], Chang-Chih Shen[2,*], Chong-Long Huang[3], Hsueh-Ting Chu[4,8], Hsiao-Yun Kuo[1], Mei-Ling Wang[1], Wen-Hua Lin[5], Yueh-Tsu King[6], Chih-Feng Chen[7], Yen-Pai Lee[7], Cheng-Yan Kao[3,8], Chaur-Chin Chen[2,8,$]

Abstract

Automatic image data analysis (*Aida*) on the genotyping polymorphic bands of microsatellite DNA requires routinely automatic image segmentation and especially translated absolute data. Computerized automatic analysis is crucial in reproducibility and artifact-proof despite of the available commercial software which is highly interactive and hence gravely time-consuming. Albeit, the genotyping data of microsatellite DNA fragment that is represented as relative mobility sizes with floating points based on selected DNA size standards may lead to great loss in data integration. The implemented *Aida* system with general-*Aida*-modules (*GAM*) of object segmentation and specific-*Aida*-modules (*SAM*) of data translation intends to provide efficient solutions in its semi-automatic format. In an experiment of twenty images, the *Aida* system demonstrated 85~95% accuracy in automatic band segmentation. Moreover, the *Aida* system along with in-house DNA size standards of *MarkQoff* (Marker-Quarter-Off) successfully translated absolute size data of DNA fragments with zero standard deviation while analyzing triplicate genotyping images of three running distances. In contrast to *Aida* system's superior performance, the results of commercial software based on fifteen DNA size standards in triplicate genotyping images delivered the DNA fragment size data up to 2.19 base pair of standard deviation along with the discrepancies of data floating points and missing bands.

Keywords: Automatic Image Dada Analysis, Object Segmentation, Absolute Size, Band, Lane, Size Standard.

[1] Breed-Use-Special Laboratory, Center of Agriculture Hierarchical Utilization, Graduate Institute of Biotechnology, College of Agriculture, Chinese Culture University, Taiwan; *, first author.
[2] Department of Computer Science, National Tsing Hua University, Taiwan; *, first author of equal contributions; $, corresponding author.
[3] Department of Computer Science and Information Engineering, National Taiwan University, Taiwan.
[4] Department of Computer and Information Engineering, Asia University, Taiwan.
[5] Animal Health Research Institute, Council of Agriculture, Executive Yen, Taiwan.
[6] Animal Technology Institute Taiwan, Taiwan.
[7] Department of Animal Science, National Chung Hsing University, Taiwan.
[8] Angiogenesis Research Center, National Taiwan University, Taiwan.

1. Introduction

Genotyped polymorphic bands on bio-images of electrophoretic analysis are represented as relative mobility sizes of DNA fragments based on selected DNA size standards. Polymorphic mobility patterns of interested band objects (*BO*) are often analyzed among individual subjects with respective DNA samples to compare genotypic relations. Importantly, relative mobility size data with floating points has imposed serious obstacles to genotyping among data integration, procedure standardization, and system automation with available commercial software packages [1~4].

The bio-images are with similar image objects produced by bio-techniques with practical purposes of genotyping, microarray and two-dimensional gel. Indeed, indispensable image data analysis software bridges the revolving bioinformatic data analysis and bio-experimental tasks. With dynamically adjustable intensity range for marking *BO* from local background, boundary drawing of *BO* is crucial for accurate genotyping image analysis [5~9].

The basic modules of object segmentation and data translation in automatic image data analysis (*Aida*) algorithm demands complete implementation to meet the practical specifications of general object segmentation and specific task goals. Segmenting *BO* of microsatellite genotyping images essentially deals with contrast quality of local background and objects. With dynamic threshold based on contrast quality, object segmentation sensitively locates object boundaries better than the cases based on simple average filter of noise elimination especially with the cases of spread-out and blur objects [10, 11].

In this paper, we implement general *Aida* modules (*GAM*) of object segmentation to reach the 85~95% accuracy in standard procedure with 20 image samples with minimal user efforts of setting 2 parameters. In addition, we implement specific *Aida* modules (*SAM*) of lane finding and marker localization to translate the absolute size data in nucleotide lengths. Specifically for genotyping tasks, *GAM* computes single band with object segmentation and *SAM* respectively computes grouped bands with lane finding and marker localization of absolute size standards to translate absolute nucleotide number of microsatellite genotyping fragments [12,13].

2. Materials and Methods

2.1: Microsatellite Genotyping Analysis

1. Microsatellite DNA Amplification

The genomic DNA samples of total 6 chickens are analyzed with ADL0102 IRDye700-primer (Tm 54.3°C) at regular PCR program, respectively. The PCR reaction is carried out under 36 [(denature), (anneal), (react)] cycles at [(°C, second)] conditions of [(95,30), (52,15), (72,30)] and initial hot start. The ADL0102 IRDye700-primer is from

custom synthesis for Primus 96 PCR Thermocycler (MWG Co., Germany).

2. Prepare Size Standards: SAGA & MarkQoff

The HPLC-purified *MarkQoff* DNA fragments (50 to 1,500 bp) of mixed dd(A,C,G)TP sequencing yield are the enriched absolute size standards based on the applied sequencing template with known MarkQoff sequence (*MQS*). The *MQS* template is sequenced with T7 IRDye700-primer for 36 [(denature), (anneal), (react)] cycles at [(°C, second)] conditions of [(95,30), (50,15), (70,30)] with initial hot start. The T7 IRDye700-primer is from custom synthesis for Primus 96 PCR Thermocycler (MWG Co., Germany) with Sequi Therm EXCEL™ II Kit (EPICENTRE Co., USA) and for D-7000 HPLC purification (Hitachi Inc., Japan) with Triethylammonium acetate buffer (Sigma Inc., USA). In addition, commercial *SAGA* size standard including 50, 75, 94, 100, 105, 120, 145, 175, 200, 204, 230, 255, 300, 325, and 350 bp is purchased thereat (LI-COR Inc., USA).

3. Image Analysis: BioNumerics and Aida Software

With 10 sample lanes in-between interval marker lanes, the denatured materials including genotyping samples and marker samples of *SAGA* and *MarkQoff* are separated on Global System 4200 auto-sequencer (LI-COR Inc., USA) with 0.6 uL material loaded, respectively. The 0.2x250-mm gel of 5.5% polymerized acrylamide gel in 1x TBE buffer with Urea is installed for electrophoresis at 45°C condition with 1200 volt, 25amperate, and 30 watt program. The gel image of IRDye700-primed DNA signals is captured as 16-bit TIFF file upon reaching gel end-zone by the laser scanner set at the speed of 4 Fast. The genotyping image data is analyzed with commercial BioNumerics software package (Applied Math Inc., USA) by setting the size standard information of both SAGA and MarkQoff for getting floating points data of relative mobility size. Moreover, the 16-bit TIFF image data is analyzed with in-house *Aida* software of which the implementation details are in the following sections.

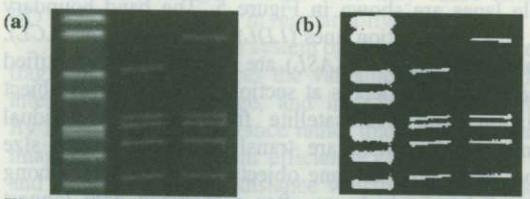


Figure 1. Locating band objects (BO) with general Aida modules (GAM). (a) Original gel image, and (b) BO acquisition.

2.2: Single Object Segmentation by GAM

Locate Band Objects (BO)

Let I(x,y), $0 \leq x < R$, $0 \leq y < C$, be an input gel image with R rows and C columns. The goal is to locate all of the band objects (BO, brighter regions) in the image based on two-way image segmentation techniques. With the input image in Figure 1, we first project the pixels onto the horizontal direction by $H(y)=\frac{1}{R}\sum_{x=0}^{R-1} I(x,y)$, $0 \leq y < C$, and separate the vertical lanes according to the valleys of H(y) distribution. Then for each lane, we apply image segmentation based on the histogram of gray levels on each individual lane. Further, overlapping lanes and bands can be solved by computing physical distances between both two neighboring lanes and two adjacent bands. Finally, the noise can be eliminated by using v*h mean or median filters. Each band object can be represented by the centroid associated with the average gray level of pixels detected in the corresponding band region.

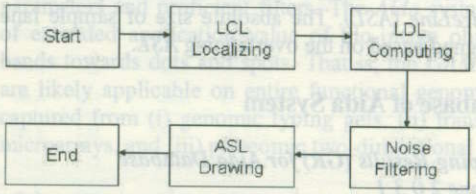


Figure 2. Flow chart of specific Aida modules (SAM).

2.3: Grouped Object Translation by SAM

1. Lane Central

The *SAM* tracks on sample lane of genotyping *BO* with one *LaneCenterLine (LCL)* and two *LaneDemarcationLines (LDL)* bilaterally. The *CenterPixel (CP)* is the center of *BO* while the band size with surrounded pixels number is between one third and 1.2-fold of *ABWH* as of normal bands. The band width and height are the absolute difference of extreme pixels respectively from x-coordinate and y-coordinate. The center position of *CP* is deleted if band width is bigger than *AverageBandWidth (ABW)*. The x/y-coordinate of *CP* is the mean of total pixels summation of both x/y-coordinates in normal *BO* to avoid mistakes in lane positions. Object segmentation is completed with normal band as for the grouped objects of respective genotyping sample.

2. Lane Tracking and Construction

A *LeastSquareLines (LSL)* of any three consecutive *CP* is drawn as of linear algebra if the variable x/y-distance of the x/y-coordinate difference of *CP* is respectively less than *ABW* or L* *AverageBandHeight (ABH)*. The L**ABH* installs an efficient dynamic threshold of x distance in practice. The *LCL* is the fused *LSL* of overlapping *LSL* pair after recursive processing of paired *LSL* until no intersected overlapping *LSL*. The algorithm is listed in the following description. Extending *LCL* is according to lane's trends in *LCL* slope in order to cross as many bands as possible aside from other lane's *LCL* crossed bands. With *LCL* linking band centers of lane, the bilateral *LDL* defines the left and right boundaries of sample lane whereas the reality left-and-right boundaries are computed as half *ABW* distance bilaterally from the *LCL* y-coordinate. To avoid reducing or broadening the various lane widths, the lane boundaries are set by *BB* x-coordinate which is closest to Pixel(n,m) as of *Neighbor.BB.X(n, m)* with n of row coordinate. With revealed *LCL* and *LDL* of respectively grouped lane objects, the similar trends of two adjacent lanes with *LCL ABW**n in y-axis direction may shift into new neighboring *LCLs* (n=-1,1,-2 or 2) without any intersecting between lanes. The *Aida* system applies *LDL* points as *BB* in order to filter off small noise bands which are not crossed by *LCL* of respective lane.

3. Marker Localization and Translation

The *Marker Center Lines (MCL)* is the size standard lane with longest *LCL* length and with greater than 0.9 of *LCL/LL* ratio. The *MCL*-crossed *BO* of size standard lane are read as *B* (*BO*) and as k* *N* (none object, *NO*) inserted at adjacent-*BB* gap with extended distance than regular unit distance. The k-integer for *N* gap of *BN*-string is determined by the ratio of gap distance over *ABH* unit distance equal to (k+1) integer. Between *BN*-strings, respective *BO/NO* are aligned with *MarkQoff Sequence (MQS)* to assign the number of corresponding absolute

size. The *MQS* is translated as *B* at ACG and as *N* at T towards *BN*-string. The *BO/NO* of identical absolute size number among marker lanes are grouped by respective lane positions and are linked-up as horizontally parallel *AbsoluteSizeLine (ASL)*. The absolute size of sample lane BO is assigned based on the overlapping *ASL*.

2.4: Database of Aida System

1. Genotyping Results (GR) for Aida Database

Marker Lane 2 0 3 1
Sample Lane 4 5 6 7 8 9
Lane 0 216 217 219 220 222 223 224 225 227 228
Lane 0 230 231 232 233 234 ... ...
PIXEL 2 1 1087 17 8189 1 1 217
PIXEL 2 1 1086 17 6930 0 1 217
PIXEL 2 1 1086 18 7968 0 1 217
PIXEL 2 2 1105 16 8957 1 1 219
PIXEL 2 2 1104 16 8538 0 1 219

The *GR* text file saves the Aida analysis results of *BO* including band number, marker location, and band intensity for loading into Aida database.

(1) MARKER LANE (*ML*), SAMPLE LANE (*SL*), Lane and PIXLE are the record types of *GR* shown at the first word of data line with TAB separator.
(2) MARKER LANE records marker lane indices.
(3) SAMPLE LANE records sample lane indices.
(4) LANE records lane index at second word along with maximally 10 size marker locations. The *ML* or *SL* requires many Lane data lines to record all the marker location of *BO*.
(5) PIXEL records lane index, band index, x-coordinate, y-coordinate, intensity, band center, marker lane, and marker location of *BO*.

2. Genotyping Analysis

By the *GR* example, 4 *ML* and 6 *SL* are shown with marker locations at 0-th lane including 216, 217, 219, 220, 222, and so forth. PIXEL [1087,17] is the center point of 1st *BO* at 2nd lane of marker lane with intensity of 8189 and absolute size of 217. For the genotyping analysis, the *GR* data is loaded into *Aida* database including sample lanes with band absolute size and band intensity for the band intensity cutoff ratio in each lane along with detailed experimental information.

3 Results and Discussion

3.1: Single Band Object Segmentation

Our *Aida* system applies local background threshold and band size respectively to filter noises of low-frequency and high-frequency automatically. The segmented object of single genotyping band is shown by the band boundaries (*BB*) of exemplified band object (*BO*) in Figure 3. The lane *BO* is efficiently located with flexible *BB* in overlapping stretches of *BR* as shown in Figure 4. Descriptive details of the algorithm are at section 2.2 of general *Aida* module (*GAM*) on single object segmentation.

Our Aida software may work more accurately with realistic genotyping bio-images of slab gel electrophoresis without knowing band regions at first of which is precisely the idealized provision in commercial image data analysis software of genotyping gel images. Specifically, our Aida system intends to resolve the issues of excessive manual interaction and accurate band segmentation by automatic processing with only two parameters of local background threshold and band size entered by user. Additionally, the peak, valley, band size, and local mean of pixels intensity are also included for automatic band object segmentation in Aida system in order to rectify the fact that these criteria are not universally efficient due to the required provision of ideally uniform band object and lane.

Commercial software for gel image often requires to manually set lane positions first before band segmentation. Moreover, some software may overlook image processing issue with only vertical waveform peaks to segment band objects. Especially, some software may mistakenly segment band objects by presumed regular shape of rectangle in the genotyping gel image.

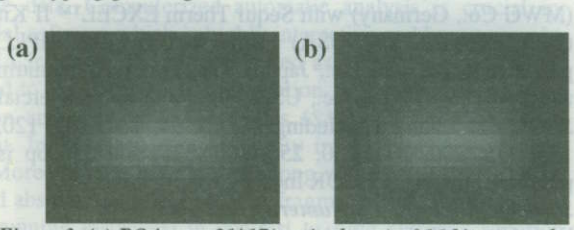


Figure 3. (a) BO image 21*17(row*column) of 16-bit grayscale. (b)BO segmentation with BB is set at red pixels. The BO and BB stands for band object and band boundary.

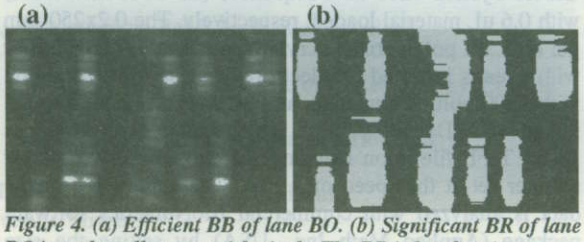


Figure 4. (a) Efficient BB of lane BO. (b) Significant BR of lane BO is set by yellow-greenish pixels. The BR is band region.

3.2: Grouped Lane Objects Translation

With specific *Aida* modules (*SAM*) for polymorphic genotyping images data analysis, band objects in respective sample lanes are shown in Figure 5. The band boundary (*BB*), lane demarcation lines (*LDL*), lane center line (*LCL*), and absolute size lines (*ASL*) are automatically identified by the descriptive details at section 2.3 of grouped object translation. The microsatellite fragments of individual genomic DNA sample are translated into absolute size number as of grouped lane objects based on ASL among *MarkQoff* lanes which are often recognized with longest image coverage.

Our *Aida* algorithm guarantees that all *LSL* across respective lane center with both *ABW*-limited lane *CP* and L**ABH*-limited straight lane. Greater L value may be equal to the number of lane objects crossed by the straighter *LCL* in lane tracking. The lane tracking methods may resolve the skew issues of object lanes and genotyping image. In addition, the *SAM* noise filter herein by both band size and lane position may install appropriate *Aida* procedure for genotyping tasks and rescue mis-deletion on meaningful weaker *BO* in which occur often at initial analysis stage without valid certification.

3.3: Microsatellite Genotyping Analysis

The Aida system and MarkQoff size standards has successfully translated the absolute size data of genotyping DNA fragments with standard deviation at zero base pair by analyzing the triplicate genotyping images of identical running distance in Li-Cor Global System 4200 (Figure 6 and Table 1). With same genotyping images, commercial

software of BioNumerics and fifteen DNA size standards of SAGA delivered the fragment size data with floating points and missing bands and the major discrepancy of standard deviation at up to 2.19 base pairs.

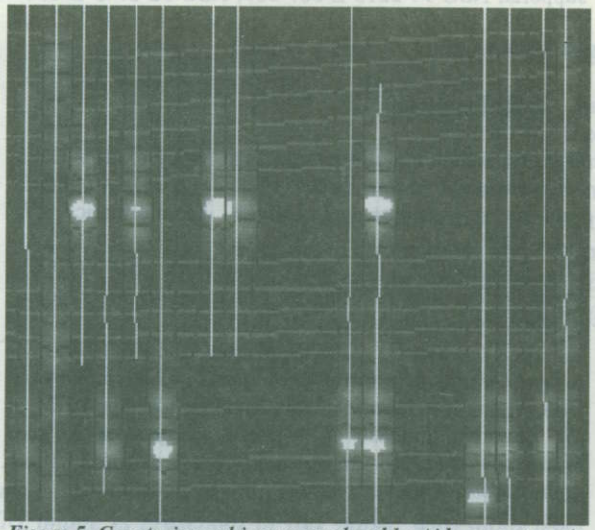


Figure 5. Genotyping gel images analyzed by Aida on automatic band segmentation, lane finding and marker translation. The BB, LDL, and LCL are set at the pixels in red, blue, and yellow curve lines. The ASL is set at the pixels in green lines. The BB, LDL, LCL, and ASL are band boundary, lane demarcation lines, lane center line, and absolute size lines, respectively.

Notably, the translated size numbers of genotyping fragments are within the small range of 112~125 base pairs which ameliorates the discrepancy by the regular software class compared herein. Likely, radiogram of genotyping on conventional sequencing gel may install additional adverse factor of variable running distance with which the mobility of size standards may be variable and affecting the relative mobility size data into greater standard deviation. Smile effect of irregular gel electric field may cause erroneous BO size data translation in regular software cases.

Overall, the Aida system of absolute size data of BO fragment length enhances the valuable data integration of inter-sample, inter-image, and inter-laboratory scenarios. By measuring band distance rather than x/y-coordinates in images, the Aida system efficiently resolves smile effect and extended mobility distance which may often result in unacceptable size data translation. The achieved accuracy of 90~100% by Aida system efficiently demonstrates its superior software performance in resolving crooked lanes and smile effect of distorted gel images.

For loading Aida database, the Aida system further sums up individual BO intensity for entire lane objects intensity towards optimizing cutoff ratio for BO acceptance in automatic and standardized processing. The translated absolute length number of accepted BO enables practical data integration of different experiments, databases, and laboratories and streamlines subsequent data analysis of specified tasks as of allelic genotyping analysis (AGA) and quantitative trait loci (QTL).

4 Conclusion

In this paper, we implement the *Aida* system of likely better performance on system automation and data integration in which may be of possibly future value for *AGA* and *QTL*. The *Aida* database carries the characteristic absolute size data of genotyping fragments and the better *BO* intensity for optimizing lane intensity percentage of the *BO* acceptance threshold towards biologically significant genotyping fragments. Thus, the specification component parts of *Aida* system include *GAM* and *SAM* with minimal parameters and proficient filters. The *Aida* system may be of extended application value of bio-image objects from bands towards dots and spots. That is, the *GAM* and *SAM* are likely applicable on entire functional genomic images captured from (i) genomic typing gels, (ii) transcriptomic microarrays, and (iii) proteomic two-dimensional gels.

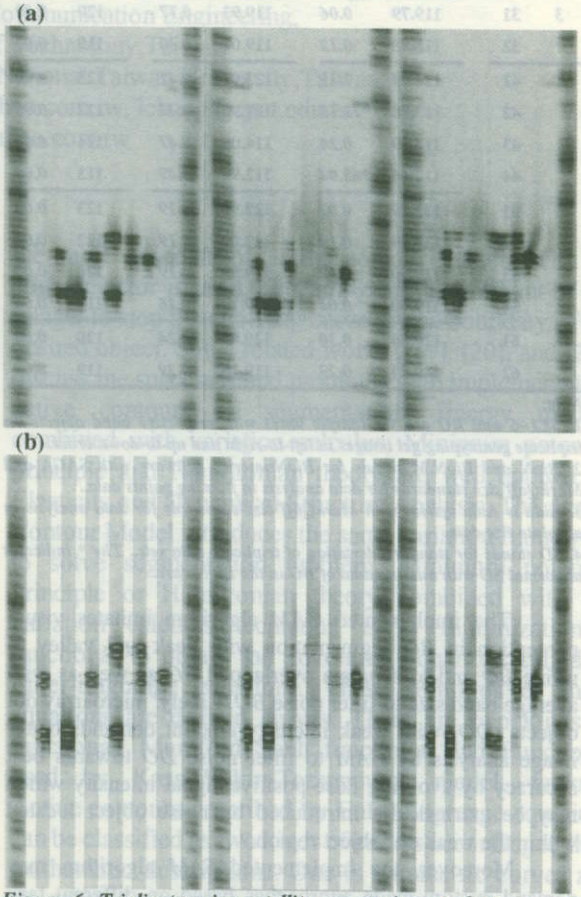


Figure 6. Triplicate microsatellite genotyping analysis on gel images of same running distances by Aida system and in-house MarkQoff DNA size standards. (a) Raw image. (b) BioNumerics software. The lanes of each sample image are M, S ,1, 2, 3, 4, 5, 6, S, and M. The M and S lanes are MarkQoff and SAGA size standards. Lanes 1~6 are six microsatellite genotyping samples of different chickens amplified with ADL0102 IRDye700-primer pair.

In general, genomic typing gel reveals polymorphic bands image of mobility patterns upon various fragments sizes by which genotypic relation of subjects are quantified statistically as genetic distance. Transcriptomic microarray reveals orthogonal chessboard dots image of hybridization signal upon various mRNA levels by which up-or-down ratio based on normal control are analyzed statistically as scenario-related feature gene. Proteomic two-dimensional gel reveals geometric chessboard spots image of staining signal upon various protein amounts by which up-or-down ratio based on normal control are analyzed statistically as scenario-related feature protein. More importantly, analysis of proteomic image combine the methods applied in both genomic and transcriptomic images.

Table 1. Translated size data by Aida system and BioNumerics software with MarkQoff and SAGA size standards.

SL	BO	BioN	SD	BioN.M	SD	Aida	SD
1	11	119.83	0.26	119.87	0.05	120	0.00
	12	118.73	0.28	118.92	0.02	119	0.00
	13	113.36	0.29	113.77	0.08	114	0.00
	14	112.45	0.39	112.72	0.00	113	0.00
2	21	113.48	0.22	113.90	0.06	114	0.00
	22	112.51	0.15	112.84	0.14	113	0.00
3	31	119.79	0.06	119.93	0.17	120	0.00
	32	118.88	0.12	119.01	0.30	119	0.00
4	41	124.09	2.19	123.02	0.42	123	0.00
	42	123.28	*71.18	121.99	0.47	122	0.00
	43	113.89	0.26	114.08	0.47	114	0.00
	44	112.66	*65.04	112.99	0.49	113	0.00
5	51	122.89	0.02	122.98	0.19	123	0.00
	52	121.99	0.19	122.07	0.19	122	0.00
	53	119.90	0.12	119.99	0.19	120	0.00
	54	119.00	0.04	119.10	0.16	119	0.00
6	61	120.04	0.30	120.07	0.24	120	0.00
	62	119.14	0.23	119.18	0.29	119	0.00

Note:
1. SL1~6 and BO are six sample lanes with respective band objects on triplicate genotyping gel images in left-to-right and up-to-down order.
2. BioN and BioN.M stands for BioNumerics software with SAGA and MarkQoff size standards for data analysis in floating points data..
3. Aida is Aida system with MarkQoff size standards for data analysis in absolute size data.
*4. SD stands for standard deviation of triplicate data sets.. The * indicates abnormal SD with missing data of BO in the BioN case.*

The implemented *GAM* algorithm imitates visual perception in *BO* segmentation with peak and valley to refine *BO* and to separate overlapped *BO* as well as with average band size to filter noise *BO*. Likely, microarray dot objects (*DO*) with weak intensity amount demands better image analysis by *GAM* to filter noise *DO* towards best accuracy by avoiding false-positive signal intensity which may be extrinsically introduced by noise object located within the weak dot object region.

Moreover, the implemented *SAM* algorithm may demand custom-made approaches based on different task goals of genomic, transcriptomic, and proteomic analysis. The common theme is to defeat non-ideal image quality as of processing difficulty towards best data translation after proficient grouping of automatic analysis. Genotypic *Aida* system achieves automatic lane finding by overcoming non-straight lines of distorted and deformed images. Yet, commercial genotyping software defining lanes by manual drawing or assigning rectangles realistically unlikely to locate lanes and objects is less efficient and accurate. Very likely, two-dimensional gel spot objects (*SO*) of composite mobility lanes demands better image analysis by *SAM* to group and match geometric chessboard *SO* as of isoelectric focusing and molecular weight separation performed in orthogonal directions towards best accuracy by avoiding false-positive global superimposing which is mistakenly accomplished by the erroneous feature *SO* selected for *SO* image superimposing between experiment and control set. With accurate superimposing, the inter-gel protein amount comparison of *SO* signal level on 2D-gel image set may be faithfully achieved for observing constituent and amount differences between experiment and control samples.

5 Acknowledgments

The authors would like to thank the National Science Council (NSC), the Council of Agriculture (COA), and the Department of Industrial Technology (DoIT) of Ministry of Economic Affairs (MOEA), Taiwan for financial supports: NSC 94-2213-E-007-089, NSC 92-2745-B-034-001, COA 93 AS-3.1.2-AD-U1, and MOEA 93-EC-17-A-19-S1-0016.

6 References

[1] Umeshadiga PS, Bhomra A, Turri MG, Nicod A, Datta SR, Jeavons P, Mott R and Flint J (2001). *Automatic analysis of agarose gel images.* Bioinformatics 17: 1084-1089.

[2] Umeshadiga PS and Chaudhuri BB (2000). *Segmentation and counting of FISH signals in confocal microscopy images.* Micron 31: 5–15.

[3] Ye XY, Suen CY, Cheriet M and Wang E (1999). *A Recent Development in Image Analysis of Electrophoresis Gels.* Vision Interface '99, Trois-Rivières, Canada, 19-21.

[5] Zhao J, Shimazu Y, Ohta K, Hayasaka R and Matsushita Y (1996). *An outstandingness oriented image segmentation and its application.* Int. Symposium on Signal Processing and its Applications, August.

[6] Tsai MY, Chang CF, Chu HT, Chan CH, Chang KJ, Kao CY and Chen CC (2005). *Microarray Images Pre-Analysis for Critical Gene Expression Computation with Implemented Algorithmic Kernel.* Hwa Kang Journal of Agriculture 16: 43-50.

[7] Appel RD, Vargas JR, Palagi PM, Walther D, and Hochstrasser DF (1997). *Melanie II - a third-generation software package for analysis of two-dimensional electrophoresis images: II. Algorithms.* Electrophoresis 18: 2735-2748.

[8] Bettens E, Scheunders P, Sijbers J, Van DD, and Moens L (1996). *Automatic segmentation and modeling of two-dimensional electrophoresis gels.* International Conference on Image Processing 2: 665-668.

[9] Umeshadiga PS and Chaudhuri BB (1999). *Efficient Cell Segmentation Tool for Confocal Microscopy Tissue Images and Quantitative Evaluation of FISH Signals.* Int. J. Microscopy Res. Technique 44: 49-68.

[10] Umeshadiga PS and Chaudhuri BB (1998). *An efficient cell segmentation tool for confocal microscopy tissue images for quantitative evaluation of FISH allele bands.* Int. J. Microscopy Res. Techniques 43: 1-20.

[11] Conradsen K and Pedersen J (1992). *Analysis of two-dimensional electrophoresis gels.* Biometrics 42: 1273–1287.

[12] Lindeberg T (1998). *Feature Detection with Automatic Scale Selection.* International Journal of Computer Vision 30: 79-116.

[13] McCarthy PJ, Sweetman SFS, McKenna PG and McKelvey M (1997). *Evaluation of manual and image analysis quantification of DNA damage in the alkaline comet assay.* Mutagenesis 12: 209-214.