Title	Imaging approaches for chromosome structures
Author(s)	Fukui, Kiichi; Kato, Seiji
Citation	Chromosome Research. 2021, 29, p. 5-17
Version Type	AM
URL	https://hdl.handle.net/11094/78960
rights	
Note	

Imaging approaches for chromosome structures

Kiichi Fukui[1*] and Seiji Kato[2]

[1.] Graduate School of Pharmaceutical Sciences, Osaka University, Yamada-oka, Suita 565-0871, Osaka, Japan

[2.]Kobe Active Aging Research Hub (KAARb), Graduate School of Human Development and Environment, Kobe University, Tsurukabuto 657-8501, Kobe, Japan

* Corresponding author, ORCID ID, 0000-0002-9156-819X

Acknowledgments:

We thank the members of the Chromosome Link for helpful comments and discussion. This work was supported by the grant to K.F. from The Japan Science and Technology Agency (JST) (Strategic International Collaborative Research Program JPMJSC17E2).

Author contribution:

KF and SK devised and designed the study. They also performed survey and analysis data. Both authors read and approved the manuscript.

Keywords: Quantitative chromosome map; Condensation pattern; CP (Condensation profile); CHIAS (Chromosome Image Analyzing System); ImageJ; Machine learning

1. Introduction

Chromosomes are an interesting research target in the biological sciences. Chromosomes displayed a structure visible under an optical microscope even in the mid-19th century when chromosomes were first discovered. Furthermore, chromosomes are carriers of genetic information, the essential code of life. The nature of chromosomes is understood when their structural and the functional aspects are integrated in a balanced manner, like two sides of a coin. The structural understanding gained visually is crucial since over 80% of all learning is acquired through visual sense. However, research into structural aspects of chromosomes is not as balanced as research on functional aspects. For example, a standard method for comparing data obtained with different imaging methods remains unavailable. Furthermore, no guidelines for basic precautions for the use of imaging methods to chromosomes are universally accepted. In this review, we begin with a historical overview of the development of imaging methods for the structural analysis of chromosomes mainly exemplifying CHIAS (Chromosome Image Analyzing System), then discuss the effective imaging methods for chromosome studies. We also cover new methods that are expected to be available in the near future.

2. Image analysis of chromosomes

a. Chromosomes are visible carriers of genetic information

Plant chromosomes are found mostly in sizes ranging from 1 to 10 μm. Typical images of small chromosomes of rice (Fig. 1 right, genome size: 430 Mbps) and large chromosomes of barley (Fig.1 left, genome size: 5,100 Mbps) are representative (Fukui et al. 2000) . Substantial variability exists in both chromosome sizes and numbers even among plant species within the same family, Gramineae. Structural information, such as the length of individual chromosomes, length of short and long arms, total number of chromosomes per cell, and presence of satellite chromosomes, are essential for a basic understanding of organisms. For example, the large chromosome rearrangement confirmed in Miyakogusa (*Lotus japonicus* L.) (Ito et al. 2000a) could explain distortion linkage maps of chromosomes 1 and 2 of varieties, Gifu B-129 and Miyakojima MG-20 (Hayashi et al. 2001). Distortion in recombination value between these varieties at the specific chromosomal regions is due to the translocation between chromosomes, reconfirmed by chromosomal image analysis and with FISH methods using 5S and 45S rDNA sequences (Ohmido et al. 2010; Ohmido et al. 2007).

b. Thirty-five years of Chromosome image analyzing system, CHIAS: A historical perspective

The imaging of chromosomes was initially introduced for the analysis of fluorescence-stained images of human chromosomes (Caspersson et al. 1971a; Caspersson et al. 1971b) and banding images (Lubs and Ledley 1973; Marimuthu et al. 1974). In addition, automated detection of metaphase chromosomes by Cytoscan 110 system (Finnon et al. 1986) and classification of chromosomes for diagnostic purposes (Lundsteen et al. 1980) were attempted. Image analysis consists of several steps, such as pre-processing, segmentation, intermediate-processing, measurement, and post-processing. Consequently, integrated chromosome image analysis systems, which can automatically proceed all these steps sequentially within a system, were developed in 1980s. These systems were mainly used for karyotype analysis of human chromosomes. For example, a stand-alone system (Magiscan Image Analizer) that automated the analysis of a large number of samples was developed to diagnose human chromosomal abnormalities (Graham and Taylor 1980). High diagnostic demand and large markets drove the development of imaging methods for human chromosomes, whereas plant chromosome analysis was limited to purely scientific fields. Still, an image analysis system for plant chromosomes (CHIAS), which could proceed all the imaging steps satisfactorily within a system, was developed based on general-purpose image analysis systems, IBAS I and II (Kontron, West Germany) (Fukui 1985; Fukui 1986). The characteristics of CHIAS included (1) adaptations for the analysis of plants, demanding substantial versatility and capacity to evaluate chromosomes varying from four to 120 in number; (2) a man–machine interface to integrate researcher decisions into an analytical process to facilitate chromosome analysis; (3) automation of routine operations; and (4) an integrated microscope with automatic focusing and scanning functions.

The specifications of CHIAS hardware included an 8 bit CPU with 64 kB of RAM and 16 kB of graphic memory. The image processing unit was equipped with a micro-programmable array processor, a video I/O board for input and output of image information captured from a TV camera integrated with a microscope or on a copy stand, and a 4 MB image memory. Image memory had a resolution of 512 × 512 pixels, and each pixel had 8 bit, 256 greyscale density levels. A major feature of CHIAS was several flexible image filters with selectable parameters, which enabled semi-automatic image processing of target images with macro-programs run sequentially. This series of macro-programs was assembled for image processing, chromosomal image capture, correction and normalization, and interactive image processing as needed. Finally, the desired image information was obtained. Representative achievements of CHIAS used included complete identification of rice chromosomes and development of a quantitative chromosomal map (Apisitwanich et al. 2000; Fukui and Iijima 1991; Fukui and Iijima 1992; Iijima et al.

1991; Sparacino et al. 2004). Eighty-one years passed from an initial count of chromosomes (Kuwada 1910) to objective identification of individual rice chromosomes. A quantitative barley N-banded chromosome map was also developed using CHIAS. The considerable distortion between chromosome and recombination maps was detected based on the quantitative map of barley chromosome 6, clearly illustrating the importance of structural information for barley genetics and breeding (Fukui and Kakeda 1990; Gustafson et al. 1990).

CHIAS II improved the original version of CHIAS including an upgrade mainframe from IBAS to VIDAS (Kontron). Built-in image memory capacity was greatly expanded (Nakayama and Fukui 1997). The CHIAS II was developed for the analysis of *Brassica nigra* chromosomes and was also used for the identification of three diploid *Brassica* chromosomes (*B. rapa* 2n = 20, *B. nigra* 2n = 16, and *B. oleracea* 2n = 18, with AA, BB, and CC genomes, respectively). CHIAS II data were used to develop quantitative chromosome maps (Fukui et al. 1998). The identification of individual *Brassica* chromosomes was achieved 76 years after Karpechenko(1922) accurately determined the chromosome numbers of each species. All the chromosomes of *B. napus* (2n = 38, AACC genome) were also identified by CHIAS II, and the quantitative chromosome map or idiogram was developed (Kamisugi et al. 1998).

CHIAS III was developed for the image analysis of chromosomes in wild sugarcane, *Saccharum spontaneum* L. (AP301, 2n = 4x = 32), with significant improvements in response to rapid advances in personal computers (MacOS) and the introduction of the Internet. The conventional stand-alone image analysis system was replaced by a PC alone, and basic image analysis software was downloaded from and used on the Internet (Kato and Fukui 1998). Image brightness was 8-bit, 256 gray levels as in CHIAS I and II but could handle grayscale images of this quality using a single PC (Kato and Fukui 1998). Software for CHIAS III was distributed to researchers free of charge as a program written in a pascal-like macro-language with NIH Image (developed at the U.S. National Institutes of Health and available on the Internet at http://rsb.info.nih.gov/nih-image/) running on a PC as the host application (http://www2.kobe-u.ac.jp/~ohmido/cl/chiasIII/index.htm).

Significant advances in image input hardware were also available. We shifted input of chromosomal images from a TV or film camera to devices for the input of digital images directly using a high-resolution charge-coupled device (CCD). Cooled CCDs that could amplify particularly weak signals became indispensable for recording weak fluorescent signals from FISH experiments. CHIAS III received digital images directly from a CCD camera system, making combinations with various input systems possible. Sugarcane (Ha et al. 1999) and spinach chromosomes (Ito et al. 2000c) were identified, and quantitative chromosome maps were developed using CHIAS III. FISH images were previously analyzed by CHIAS (Fukui et al. 1994), and images obtained with CHIAS III by multi-color FISH were analyzed and supported using separated RGB channels, called “Stack” which assigns each of the three grayscale images to an RGB channel. Chromosome images after double staining with propidium iodide (PI)/4,6-diamidino-2-phenylindole (DAPI) could be analyzed by CHIAS III (Ohmido et al. 1998). A high-resolution map of rice pachytene chromosome 9 was developed using this method (Kato et al. 2003).

CHIAS IV, a current version, is compatible with both Windows and MacOS and supports 24-bit RGB colors (Kato et al. 2009). The host application is ImageJ, a Java application (Schneider et al. 2012). CHIAS IV enables chromosomal image analysis by running a plug-in program within the application. The plug-in can be obtained from the CHIAS download site of Kobe University (http://www2.kobe-u.ac.jp/~ohmido/index03.htm). Chromosomes can be analyzed using ordinary PCs regardless of the operating system. The program has been distributed free of charge to 18 countries. ImageJ runs on Java software, and CHIAS IV does not support iPad OS and Android OS that do not support Java.

ImageJ can be expanded by adding Java programs in plug-in format. The “Straighten” program for chromosomes and the “Hyperstack” format, described in Section 4.a. and 4.b., respectively, are used as commands in ImageJ. Additionally, “Fiji” (https://fiji.sc) was developed as a software package to enable easier use of many plug-ins. The file reading function of the machine learning application, “Ilastik” (Sommer et al. 2011), described in Section 4.d., is installed as a plug-in, and “Fiji” can easily link with “Ilastik.” Furthermore, “µManager” (https://micro-manager.org) is specialized for microscope operation. Both “Fiji” and “µManager” can be used for interactive karyotype analysis using CHIAS IV by installation as plug-ins. Prometaphase chromosomes of red clover were first analyzed using CHIAS IV (Kataoka et al. 2012).

c. Various chromosomal image processing and analysis technologies

Currently, DRWID (Kirov et al. 2017), which generates idiograms from Java-based FISH/GISH images, and LEVAN (Sakamoto and Zacaro 2009), an imageJ plug-in that specializes in measuring arm ratio and chromosome length, have been developed. In addition, ChIPS-Karyo (GenDix, Inc. Seoul, Korea), which is a commercial application, and Chromawizard (Auer et al. 2018), which can produce karyograms of multicolor FISH images, are known. There are libraries with various functions necessary for processing images and videos on a computer at high speed, such as OpenCV (Open Source Computer Vision Library: https://opencv.org). Although you need to have some knowledge on computer languages to use this library, an application for chromosomal image analysis using OpenCV has been developed. Chromawizard, which was introduced in Section 4.d., is a typical application. Furthermore, it supports machine learning and it is a suitable environment for automatic processing.

A comparison of Chromawizard and CHIAS IV has the following features. CHIAS IV has been developed for the analysis of plant chromosomes, whereas Chromawizard has been developed primarily for the analysis of human and animal chromosomes. Moreover, Chromawizard has features that are not available in CHIAS IV; for example, the treatment of intersecting chromosomes. When two chromosomes intersect each other, retaining one as one chromosome leaves the other as two separate objects. The two chromosomal fragments can be interactively reconnected. Such processing is not performed in the CHIAS IV Plug-in, so it is necessary to directly operate ImageJ, which is the host application. On the other hand, CHIAS IV supports karyotype analysis based on condensation patterns, which are more or less specific to plant small chromosomes. Pseudo-color display that makes it easier to view shades and color tones improvement by applying “Hyperstack” are CHIAS IV’s unique functions that Chromawizard does not have.

3. Chromosome identification and map development

a. Quantitative measurement of chromosomal structural information

The discrimination of a certain chromosome from the others or separate identification of all chromosomes has been done by using chromosome structural parameters. For example, lengths of individual chromosomes are the most easily attainable structural information. Furthermore, structural information called the arm ratio, which is ordinarily, the ratio of the length of the long arm divided by the length of the short arm is used to be employed for discrimination and identification of chromosomes. Levan et al. (1964) defined chromosomes with an arm ratio of 1.0 as mid-metacentrics, 1.0–1.7 as metacentrics, 1.7–3.0 as sub-metacentrics, 3.0–7.0 as sub-telocentrics, and 7.0 or higher as telocentrics. Such identification has been widely used, although it is sometimes difficult to identify all the chromosomes within a complement. Generally, chromosome banding methods are effective in the discrimination and identification of large plant chromosomes (Fukui and Kamisugi 1995). As a result, complete identification and development of chromosome maps for plant species with small chromosomes have lagged behind.

b. Application of imaging methods for the analyses of banded chromosomes: Overcoming the difference between machine and human visual perception

N- and C-banding for large plant chromosomes, but not G-banding, are applicable for the discrimination and identification of individual chromosomes (Kakeda et al. 1991; Kakeda et al. 1990). Many plant chromosomes were identified by banding patterns; however, the identification of band positive regions between imaging methods and human visual inspection is still not completely resolved. The identification of both large and small bands within a chromosome using thresholding with a single value of brightness is frequently not possible because of a discrepancy in the identification of band regions between visual and machine determinations. Human vision identifies band positive regions by differences in brightness levels between banded regions and surrounding regions. Thus, an imaging method to simulate human vision to identify the positive band regions regardless of sizes was developed (Fukui and Kamisugi 1995). Dynamic curve fitting gave exact results by sliding discrimination. Furthermore, the application of low-pass filters, with adequate filter matrix size, generated a reference image that could be used for this discrimination. Relevant band regions were identified by the subtraction of the filtered image from the original. Band regions identified by imaging methods thus corresponded exactly to band regions recognized visually.

c. Identification of plant small chromosomes using condensation patterns

Condensation pattern method was developed to assist with the characterization of plant small chromosomes. Condensation pattern, the differentiation between condensed and non-condensed regions especially in plant small chromosomes at prometaphase, were widely known(Fukui 1986). Although objective and quantitative analyses of chromosome condensation were not immediately available, application of imaging methods was reported using *Atriplex* chromosomes (Fukui and Mukai 1988). Rice chromosomes with small sizes (Fig. 1 right) were also known to show a shade distribution (Fig. 2a, Khush and Kinoshita 1991). The distribution of shading is not due to differences in pigment affinity. Instead, shading reflects differences in the degree of chromatin fiber condensation. Dark regions indicate a high degree of chromatin fiber condensation (Fukui and Iijima 1991). In many cases, condensation patterns can be observed with good reproducibility by omitting pretreatment (low-temperature treatment; chemical treatment, such as colchicine) that promotes chromosome condensation. Two rice chromosomes, 7 (over) and 6 (under), are enlarged and pseudo-colored in Fig. 2b. Imaging procedures enhance the visibility of detailed characteristics of condensation pattern by the human eye much more than small grayscale images. Consequently, 360 rice chromosomes from 30 haploid chromosomal spreads were completely identified by computer-aided inspection using enlargement, pseudo-coloration, and numerical data of arm ratios and lengths. CP (condensation profile) or a density profile of condensation patterns averaged by 30 chromosome 7s is shown in Fig. 2c. Heavy condensation at the proximal region of the chromosome is demonstrated. Fukui and Iijima (1991) and Iijima et al. (1991) used condensation pattern, its CP, density profile at the mid-rib of each chromatid, and numerical data of chromosome length and arm ratio to discriminate a certain chromosome from the other and identify separately all rice chromosomes. Furthermore, thresholding the averaged CP with two brightness values to represent borders between condensed and heavily condensed regions, and dispersed and condensed regions, enabled the total length of the CP to be divided into the three different regions of dispersed, condensed and heavily condensed regions, and a quantitative chromosome map for all the rice chromosomes was developed (Fig. 2c).

d. Automatic identification of the chromosome number using condensation profile, CP

Development of a computer-aided method for automatic identification of rice chromosomes was explored based on image parameters obtained by image analysis. Numerical data from the 360 condensation profiles (CPs) obtained from 360 rice chromosomes of 30 chromosome spreads of haploid rice were subjected to three different discrimination methods to explore the possibility of automatic identification of rice chromosome. The first one was the method using the discrimination flow chart. The second and third ones were the method using linear discrimination functions and the minimum distance classifier based on standardized Euclidean distance. As a result, discrimination based on the minimum distance classifier resulted in the correct identification of rice chromosome number more than 92% of the rice chromosomes. The first and second methods gave a correct identification of 91 and 84%, respectively (Kamisugi et al. 1993). It is concluded that most of rice chromosomes could be identified by a computer using numerical data obtained from the CP. It also means that the image parameter, CP is a reliable and reproducible parameter to represent individual rice chromosome structure.

e. Development of quantitative chromosome maps using condensation profile, CP

CP as described in Sections 3.c. and 3.d. enabled the development of a quantitative chromosome map for sugarcane by CHIAS III (Ha et al. 1999; Kato and Fukui 1998). Fig. 3 shows a series of steps for processing images. Wild-type sugarcane is octoploid (2n = 8x = 64, x = 8) and we used a tetraploid sugarcane (AP301, 2n = 4x = 32, x = 8) with four homologous chromosomes in each cell (Fig. 3a). A

FISH image of rDNAs with 5S (blue) and 45S (red) rDNA signals confirmed the tetraploid status of the genomes (Fig. 3b). For image analysis, the image size was first adjusted to an appropriate size for analysis of condensed regions of chromosomes (Fig. 3c), and the threshold for appropriately extracting chromosome regions from the background was interactively determined (Fig. 3d). Next, the brightness of chromosome regions was normalized to 1-255 (Fig. 3e). Next, suitable LUT was selected and applied for a pseudo-color display to facilitate visualization of difference in brightness. The visual confirmation of characteristics of the pseudo-colored image was accompanied by labeling (grouping) individual chromosomes into categories given alphabetic designations from A to H (Fig. 3f, g). Letters separated chromosomes into homologous groups based on condensation patterns (Fig. 3f). Homologous chromosomes were sorted into the same image frame, facilitating comparison of characteristics in detail (Fig. 3g). Chromosome number was finally determined based on chromosome length. When necessary, CP was checked as shown in Fig. 3h. Finally, the accuracy of identification of homologous chromosomes by combining structural and FISH data (Fig. 3b).

An idiogram was developed by averaging 56 CPs from 28 homologous chromosomes from seven chromosome complements in which all chromosomes were discriminated and identified. The idiogram of wild sugarcane based on the CPs was obtained (Fig. 4). Chromosomes are numbered in the order from longest, with the short arm (p) on top and the long arm (q) at the bottom. Chromosomes are arranged based on the position of the centromere (CEN), and the secondary constriction is indicated by an arrowhead. The relative degree of chromosome condensation is represented by highly condensed regions (black), condensed regions (gray), and non-condensed or dispersed regions (white) of the idiogram. Averaged CP was used to determine these regions. First, the brightness that discriminates the condensed and non-condensed regions of chromosomes within the complement is interactively determined by the aid of CHIAS III. The brightness value for heavily condensed regions is then determined interactively.

Quantitative chromosome maps based on CP are basic genetic information used for various purposes. Particularly, this method is useful for the development of chromosome maps in plants with small chromosomes because no banding method exists that is comparable with G banding for animal cells. Quantitative chromosome maps have already been developed with *Arabidopsis* (Ito et al. 2000b), spinach (Ito et al. 2000c), *Brassica* species (Fukui et al. 1998; Kamisugi et al. 1998), *Lotus japonicus* (Ito et al. 2000a), and rice (Fukui and Iijima 1991; Iijima et al. 1991). Chromosome maps based either on banding or condensation profiles are not only useful for a bird's-eye view of the location of specific genes or DNA sequences within the genome (Kamisugi et al. 1993) but also the map can evaluate recombination fraction. Recombination value, which is essential for genetics and breeding varies greatly depending on the region of the chromosome. Suppression of the recombination value at proximal regions and elevation at the end of chromosomes have been reported (Fukui and Kakeda 1990; Gustafson et al. 1990; Künzel et al. 2000).

4. Various imaging approaches to chromosomal structure

a. Analysis of pachytene chromosomes using the “Straighten” method

Shapes of chromosomes are not fixed and often display bends or overlaps, making image analysis difficult. This problem is especially noticeable in the early stage of mitosis and meiotic pachytene stage. Fig. 5a shows rice pachytene chromosome 9 double-stained with PI and DAPI. The chromomere, characteristic of pachytene chromosomes, is clearly observed; however, it is not easy to distinguish individual chromosomes because of their long length and overlap. The nucleolar organizing region (NOR) of chromosome 9 is specifically stained with PI and can be visually identified (Kato et al. 2003). A chromosome map was developed based on the procedure shown in Fig. 5 to make the distribution of chromomeres of the identified chromosome clear. First, a chromomere diagram was developed from the fluorescence profile at the central axis of the chromosome (Figs. 5b-f), and smoothing was performed in the lateral direction to simulate shading and swelling of chromomeres (Fig. 5h). A two-dimensional chromosome map was developed based on these images (Figs. 5k, l). It is also possible to straighten curved and overlapping pachytene chromosomal images using the “Straighten” function added to ImageJ. DAPI images of all rice pachytene chromosomes arranged in numerical order by their chromosome numbers. The position and fluorescence intensity of the chromomere were then easily compared, and the characteristics of each chromosome are also objectively compared.

By comparing different chromosome maps, an integrated chromosomal map is presented (Fig. 5n). Cytological chromosomal information could be compared with linkage and physical maps (Ohmido et al. 2018). This integrated map is basic information for future rice genetics and breeding. Particularly, integration with pachytene chromosomal maps is important for the enhancement of accuracy of low-resolution somatic chromosomal maps. “Straighten” treatment is effective for the analysis of extended DNA fiber FISH (Fig. 6, Ohmido et al. 2001) and is included as one function of CHIAS IV (Ohmido et al. 2016).

b. Improvement of chromosome visibility after FISH using “Hyperstack” in counterstain

Imaging for analysis of multi-color FISH results enables the display of brightness data of each RGB channel at the same time. However, some situations exist where RGB colorized images are more difficult to analyze than grayscale images. For example, DAPI, a fluorescent dye, is often used as a counterstain because of its affinity for AT pairs of DNA. DAPI color is processed based on the blue channel; however, some cases are difficult to evaluate using differences in brightness gradation with the naked eye (Fig. 7a). ImageJ includes an image type called “Stack,” and an improved function called “Hyperstack.” “Hyperstack” is a multidimensional image that extends the image stack to four dimensions (4D) of x (width), y (height), z (slices), and c (channel or wavelength). The dimensions can also be extended to five dimensions by including t (time frame).

LUT of B could be set to grayscale to improve visibility of DAPI images (Fig. 7a) and maintain original LUTs of R and G showing FISH signals. An RGB three-color image is integrated and color signals are displayed on chromosomes counterstained by grayscale (Fig. 7b). When a slightly bluish LUT is used, the integrated FISH image on a screen, monitor, or printed page becomes much improved in terms of human visualization. Several researchers performed similar processing with photo-retouching software, such as Photoshop; however, since these platforms are not developed for data analysis, original data (measured values) would be changed. Conversely, the advantage of “Hyperstack” is that only visibility is improved while maintaining original information in RGB channels. Processing display in images in Figs. 7b and c demonstrate only a change of LUT, and the original luminance data are maintained. Thus, the display could show the original image (Fig. 7a). This feature has been added to CHIAS IV and is available by download.

c. Three-dimensional structural analysis of human chromosomes by reconstructed 3D images

Sequential slicing of an isolated human metaphase chromosome by focused ion beam (FIB) and imaging of cross-sections of chromosomes by SEM (FIB/SEM; Focused Ion Beam/Scanning Electron Microscope) is a powerful tool (Fukui 2016; Hamano et al. 2014; Poonperm et al. 2015; Schroeder-Reiter et al. 2009; Wako et al. 2020). ImageJ 1.51 and Volume Viewer (V2.01) plug-ins were used to reconstruct 3D chromosome images from sequential serial section images of the human metaphase chromosomes (1280 × 533 pixels, 224 cross-sections). The 3D image can be rotated in any direction to view the overall surface structure using a movie (Fig. 8a, Suppl. Video1a, (Wako et al. 2020). Furthermore, the reconstructed 3D image can be virtually cut along any plane, position, or angle to view the inner structure (Fig. 8b, Suppl. Video1b, (Wako et al. 2020).

d. Application of machine learning to chromosomal image analysis

In CHIAS I-III, the chromosomal region was extracted from the background by thresholding using a moving average from the original image. In this case, the result depends much on the kernel size used to calculate the moving average. The current CHIAS IV also uses moving average for thresholding images but with different kernel sizes at the same time. In addition to this, CHIAS IV can use ImageJ's original command, background removal technology. This would facilitate the optimum background removal. Nowadays, machine learning methods are attracting attention for background removal.

”Ilastik” is an application that can be easily linked with CHIAS IV. “Ilastik” enables contour extraction without any knowledge or experience of the effects of filter kernel size (Sommer et al. 2011). “Ilastik” uses machine learning for image classification and segmentation. It is highly effective in extracting satellite regions, which was difficult with ordinary low-pass filtering (Fig. 9). “Ilastik” runs on Windows and MacOS, and since it can import files with “Fiji”, it also supports CHIAS IV. OpenCV introduced in Section 2.c. is also attracting attention when introducing machine learning. OpenCV has a rich library related to high speed image processing including machine learning, and is suitable for automatic processing.

5. Conclusion and future perspectives

Machine learning and AI will be of great help to more researchers in making decisions and choosing the right parameters in advanced chromosomal image analysis. In the near future, it will be possible to identify the number of individual chromosomes using a computer that machine-learns chromosome images with condensation patterns of plant small chromosomes or with banding patterns of plant large chromosomes and animal chromosomes. Furthermore, it will be possible to identify two homologous chromosomes or sex chromosomes in a chromosome spread by AI. Then, the method will be directly applied to automatic detection of trisomy, even smaller chromosomal abnormalities or the other diagnostically important information or both. Of course, it also contributes to basic research in chromosome science. For example, by introducing machine learning based-AI, it is anticipated that the structural features of the condensation pattern, which is perceived as uneven condensation, is not only structurally elucidated at the chromatin level using advanced microscopy, but also its function is understood in an integrated manner as one of the unique properties of chromosomes. By continuing the integrated elucidation of structure and function to the understanding of the characteristics of chromosomes, chromosomes will ultimately give the whole picture to us.

References

Apisitwanich S, Shishido R, Akiyama Y, Fukui K (2000) Quantitative chromosome map of a representative indica rice Euphytica 116:161-166

Auer N, Hrdina A, Hiremath C, Vcelar S, Baumann M, Borth N, Jadhav V (2018) ChromaWizard: An open source image analysis software for multicolor fluorescence in situ hybridization analysis Cytometry A 93:749-754 doi:10.1002/cyto.a.23505

Caspersson T, Lomakka G, Møller A (1971a) Computerized chromosome identification by aid of the quinacrine mustard fluorescence technique Hereditas 67:103-109

Caspersson T, Lomakka G, Zech L (1971b) The 24 fluorescence patterns of the human metaphase chromosomes—distinguishing characters and variability Hereditas 67:89-102

Finnon P, Lloyd D, Edwards A (1986) An assessment of the metaphase finding capability of the Cytoscan 110 Mutation Research/Environmental Mutagenesis and Related Subjects 164:101-108

Fukui K (1985) Identification of plant chromosome by image analysis method The Cell (Tokyo) 17:145-149.

Fukui K (1986) Standardization of karyotyping plant chromosomes by a newly developed chromosome image analyzing system (CHIAS). Theor Appl Genet 72:27-32

Fukui K (2016) Contribution of nanotechnology to chromosome science Chromosome Science 19:51-56

Fukui K, Iijima K (1991) Somatic chromosome map of rice by imaging methods. Theor Appl Genet 81:589-596

Fukui K, Iijima K (1992) Manual on rice chromosomes 2nd ver. Misc Pub Natl Inst Agrobiol Resour 4:1-25

Fukui K, Kakeda K (1990) Quantitative karyotyping of barley chromosomes by image analysis methods. Genome 33:450-458

Fukui K, Kamisugi Y (1995) Mapping of C-banded *Crepis* chromosomes by imaging methods Chromosome Res 3:79-86

Fukui K, Mukai Y (1988) Condensation pattern as a new image parameter for identification of small chromosomes in plants. Jpn J Genet 63:359-366

Fukui K, Nakayama S, Ohmido N, Yoshiaki H, Yamabe M (1998) Quantitative karyotyping of three diploid

Brassica species by imaging methods and localization of 45s rDNA loci on the identified chromosomes. Theor Appl Genet 96:325-330

Fukui K, Ohmido N, Khush G (1994) Variability in rDNA loci in the genus *Oryza* detected through fluorescence *in situ* hybridization Theor Appl Genet 87:893-899

Fukui K, Ohmido N, Wako T (2000) Smallness: gain and loss in plant chromosome research. In: Chromosomes Today 13. Springer Birkhäuser, pp 287-301

Graham J, Taylor CJ (1980) Automated chromosome analysis using the Magiscan Image Analyser Anal Quant Cytol 2:237-242

Gustafson J, Butler E, McIntyre C (1990) Physical mapping of a low-copy DNA sequence in rye (Secale cereale L.) Proceedings of the National Academy of Sciences 87:1899-1902

Ha S, Moore PH, Heinz D, Kato S, Ohmido N, Fukui K (1999) Quantitative chromosome map of the polyploid *Saccharum spontaneum* by multicolor fluorescence *in situ* hybridization and imaging methods Plant Mol Biol 39:1165-1173

Hamano T et al. (2014) Chromosome interior observation by focused ion beam/scanning electron microscopy (FIB/SEM) using ionic liquid technique Microscopy and Microanalysis 20:1340-1347

Hayashi M et al. (2001) Construction of a genetic linkage map of the model legume Lotus japonicus using an intraspecific F2 population DNA Res 8:301-310

Iijima K, Kakeda K, Fukui K (1991) Identification and characterization of somatic rice chromosomes by imaging methods. Theor Appl Genet 81:597-605

Ito M et al. (2000a) Genome and chromosome dimensions of *Lotus japonicus* J Plant Res 113:435-442

Ito M, Ohmido N, Akiyama Y, Fukui K (2000b) Quantitative chromosome map of *Arabidopsis thaliana* L. by imaging methods Cytologia 65:325-331

Ito M, Ohmido N, Akiyama Y, Fukui K, Koba T (2000c) Characterization of spinach chromosomes by condensation patterns and physical mapping of 5S and 45S rDNAs by FISH J Am Soc Hort Sci 125:59-62

Kakeda K, Fukui K, Yamagata H (1991) Heterochromatic differentiation in barley chromosomes revealed by C- and N-banding techniques Theor Appl Genet 81:144-150 doi:10.1007/BF00215715

Kakeda K, Yamagata H, Fukui K, Ohno M, Wei Z, Zhu E (1990) High resolution bands in maize chromosomes by G-banding methods Theor Appl Genet 80:265-272

Kamisugi Y, Furuya N, Iijima K, Fukui K (1993) Computer-aided automatic identification of rice chromosomes by image parameters. Chromosome Res 1:189-196

Kamisugi Y, Nakayama S, O'Neil CM, Mathias RJ, Trick M, Fukui K (1998) Visualization of the Brassica self-incompatibility S-locus on identified oilseed rape chromosomes Plant Mol Biol 38:1081-1087

Karpechenko G (1922) The number of chromosomes and genetic correlation of cultivated Cruciferae Bull Appl Bot Gen Plant Breed 13:3-14

Kataoka R, Hara M, Kato S, Isobe S, Sato S, Tabata S, Ohmido N (2012) Integration of linkage and chromosome maps of red clover (*Trifolium pratense* L.) Cytogenet Genome Res 137:60-69 doi:10.1159/000339509

Kato S, Fukui K (1998) Condensation pattern (CP) analysis of plant chromosomes by an improved chromosome image analysing system, CHIAS III Chromosome Res 6:473-479

Kato S, Ohmido N, Fukui K (2003) Development of a quantitative pachytene chromosome map in *Oryza sativa* by imaging methods Genes Genet Syst 78:155-161

Kato S, Ohmido N, Hara M, Kataoka R, Fukui K (2009) Image analysis of small plant chromosomes by using an improved system, CHIAS IV Chromosome Science 12:43-50

Khush GS, Kinoshita T (1991) Rice karyotype, marker genes, and linkage groups. In: Khush G, Toenniessen, GH (ed) Rice biotechnology. vol 6. CAB International and International Rice Research Institute, Wallingford, pp 83-108

Kirov I, Khrustaleva L, Laere KV, Soloviev A, Sofie M, Romanov D, Fesenko I (2017) DRAWID: user-friendly java software for chromosome measurements and idiogram drawing Comparative cytogenetics 11:747-757 doi:10.3897/CompCytogen.v11i4.20830

Künzel G, Korzun L, Meister A (2000) Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints Genetics 154:397-412

Kuwada Y (1910) A cytological study of *Oryza sativa* L Bot Mag (Tokyo) 24:267-280

Levan A, Fredga K, Sandberg AA (1964) Nomenclature for centromeric position on chromosomes Hereditas 52:201-220

Lubs H, Ledley R Automated analysis of differentially stained human chromosomes. In: Nobel Symp, 1973. p 61

Lundsteen C, Bjerregaard B, Granum E, Philip J, Philip K (1980) Automatic chromosome analysis: I. A simple method for classification of B-and D-group chromosomes represented by band transition sequences Clinical genetics 17:183-190

Marimuthu KM, Selles WD, Neurath PW (1974) Computer analysis of Giemsa banding patterns and automatic classification of human chromosomes Am J Hum Genet 26:369-377

Nakayama S, Fukui K (1997) Quantitative chromosome mapping of small plant chromosomes by improved imaging on CHIAS II. Genes Genet Syst 72:35-40

Ohmido N, Akiyama Y, Fukui K (1998) Physical mapping of unique nucleotide sequences on identified rice chromosomes. Plant Mol Biol 38:1043-1052

Ohmido N, Ishimaru A, Kato S, Sato S, Tabata S, Hayashi M, Fukui K (2010) Integration of cytogenetic and genetic linkage maps of *Lotus japonicus*, a model plant for the legume Chromosome Res 18:287-299

Ohmido N, Iwata A, Kato S, Wako T, Fukui K (2018) Development of a quantitative pachytene chromosome map and its unification with somatic chromosome and linkage maps of rice (*Oryza sativa* L.) PLoS One 13:e0195710 doi:10.1371/journal.pone.0195710

Ohmido N, Kijima K, Ashikawa I, de Jong JH, Fukui K (2001) Visualization of the terminal structure of rice chromosomes 6 and 12 with multicolor FISH to chromosomes and extended DNA fibers Plant Mol Biol 47:413-421

Ohmido N, Sato S, Tabata S, Fukui K (2007) Chromosome maps of legumes Chromosome Res 15:97-103 doi:10.1007/s10577-006-1109-7

Ohmido N, Wako T, Kato S, Fukui K (2016) Image Analysis of DNA Fiber and Nucleus in Plants. In: Chromosome and Genomic Engineering in Plants. Springer, pp 171-180

Poonperm R, Takata H, Hamano T, Matsuda A, Uchiyama S, Hiraoka Y, Fukui K (2015) Chromosome scaffold is a double-stranded assembly of scaffold proteins Scientific reports 5:1-10

Sakamoto Y, Zacaro A (2009) LEVAN, an IMAGEJ plugin for morphological cytogenetic analysis of mitotic and meiotic chromosomes. Initial version. An open source Java plugin distributed over the Internet from http://rsbweb.nih.gov/ij/ Accessed 16 January 2021.

Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis Nat Methods 9:671-675

Schroeder-Reiter E, Pérez-Willard F, Zeile U, Wanner G (2009) Focused ion beam (FIB) combined with high resolution scanning electron microscopy: a promising tool for 3D analysis of chromosome architecture J Struct Biol 165:97-106

Sommer C, Straehle C, Koethe U, Hamprecht FA Ilastik: Interactive learning and segmentation toolkit. In: 2011 IEEE international symposium on biomedical imaging: From nano to macro, 2011. IEEE, pp 230-233

Sparacino A, Halfer C, Strada E, Tano F, Ditto D (2004) Identification and characterization of somatic chromosomes of red rice (*Oryza sativa* subs. japonica var. sylvatica) by means of a computerized chromosome image method (CHIA-EA) J Genet Breed 58:295

Wako T et al. (2020) Human metaphase chromosome consists of randomly arranged chromatin fibres with up to 30-nm diameter Sci Rep 10:8948 doi:10.1038/s41598-020-65842-z

Figure legends

Fig. 1 Typical large and small plant chromosomes. Left: Barley chromosomes (2n = 14). Right: Rice chromosomes (2n = 24). Bar indicates 10 µm (Fukui et al. 2000).

Fig. 2 a. Somatic prometaphase chromosomes from a haploid rice plant (2n = 12) depicting a typical condensation pattern, uneven condensation of chromosomes. Unified chromosome numbers are given. As a result, chromosome numbers are not in order of chromosome length (Khush and Kinoshita 1991). b. Enlarged and pseudo-colored rice chromosomes depicting clear condensation patterns (Over: #7 and Under: #6). c. Condensation profile (CP), a one-dimensional representation of condensation pattern of rice chromosome 7. The distribution of brightness was measured at the axis of each chromatid. X and Y axes show relative brightness values and the number of pixels covering the entire chromosome, respectively. An idiogram based on the CP was developed.

Fig. 3 Representative steps in chromosome image analysis of tetraploid sugarcane (AP301, 2n = 4x = 32, x = 8) based on condensation profiles. a. Giemsa-stained somatic chromosomes depict condensation patterns. b. Chromosome images after multicolor fluorescence *in situ* hybridization with 45S (red signals indicated by four solid white arrowheads) and 5S rDNA (green signals by four open white arrowheads) probes. Chromosomes were counterstained with DAPI. c. Digitized and enlarged image for analysis. d. and e. Extraction of chromosomal regions by thresholding using adequate brightness. f. Application of a suitable LUT to enhance detailed characteristics of condensation patterns of individual chromosomes. g. Grouping of individual chromosomes to frame homologous chromosomes. h. Decision of chromosome number by detailed comparison and examination using graphic and numerical presentation of CP obtained from the condensation pattern. Bar shows 5 µm (Kato and Fukui 1998, Ha et al. 1999).

Fig. 4 Idiogram of sugarcane chromosomes developed by imaging methods and aligned at centromeres. Black, gray and white regions represent different degrees of condensation observed and serve as areal visual landmarks of chromosomes. 45S and 5S rDNA loci were physically mapped to specific chromosomal regions of 3p3.1 (solid doublet black circles) and 6q1.3 (open doublet black circles), respectively (Ha et al. 1999).

Fig. 5 Representative steps for the development of a quantitative rice pachytene chromosome map (a-l, Kato et al. 2003). Two images of chromosome 9 differentially stained with PI and DAPI were used to develop two chromosome maps showing heavily condensed regions within each chromomere. (k) Chromosome map based on brightness and each chromomere, (l) Index chromosome maps were developed through three major imaging steps. (m) Individual FISH chromosome images using the BAC/PAC clones, tandem repeats (Trs-A), and rRNA genes specific to individual rice straightened pachytene chromosomes. Bar shows 5 μm. (n) Integration of three rice maps; from left to right: nucleotide numbers represented by the length of the line, linkage map, meiotic pachytene chromosome map, and mitotic prometaphase chromosome map. Length ratios among 12 chromosomes were adjusted to ratios of pachytene chromosomes. Bar shows 1% region for all maps. Designation of the chromosome number and short/long arms of somatic chromosomes follows the IRGSP-1.0 database32. Chromosome number and assignment of long and short arms sometimes does not follow actual length order (Ohmido et al. 2018).

Fig. 6 a. An Extended DNA fiber (EDF) FISH image (Ohmido et al. 2001). Red (Trs-A) and green (Telomere) signals are straightened using the “Straighten” function. b. Straightened EDF-FISH images. Bar shows 5 μm.

Fig. 7 Effects of different LUT in the application, “Hyperstack.” a. An RGB image (Kato et al. 2009). b. A grayscale LUT is applied to the blue channel. c. Use of a suitable custom LUT in the blue channel to enhance visibility of FISH signals. Bar shows 5 μm.

Fig. 8 Three-dimensional reconstruction images of a human chromosome based on 224 cross-sectioned images obtained by FIB/SEM (Focused Ion Beam/Scanning Electron Microscope) . a. 3D reconstruction of human chromosome images. The 3D image can be rotated in any direction to view overall structure. b.

The inner structure is also observed by virtual cuts of the 3D reconstructed chromosome image with a plane at any position and any angle (Wako et al. 2020).

Fig. 9 Extraction of the chromosome regions of red clover by “Ilastik.” a. A FISH image. b. Image counterstained with DAPI. c. Chromosome regions extracted by “Ilastik.” Satellite regions are extracted with chromosome region. d. Chromosome regions extracted by CHIAS IV without “Ilastik.” Satellite and chromosome regions are separately extracted. Bar shows 5 µm.

Suppl. Video1 Movies for a reconstructed human chromosome based on 224 cross-sectioned images obtained by FIB/SEM (Wako et al. 2020). a. For the observation of the surface of a human chromosome, the 3D chromosome image can be rotated in any direction. b. For the observation of the chromosome interior by virtual cuts of the 3D chromosome image with a plane at any position and any angle.