Submission of coursework for Physics and Astronomy course PHASG199/299

Please sign, date and return this form with your coursework by the **specified deadline** to the Room E2.

DECLARATION OF OWNERSHIP

- I confirm that I have read and understood the guidelines on plagiarism, that I understand the meaning of plagiarism and that I may be penalised for submitting work that has been plagiarised.

- I confirm that all work will also be submitted electronically and that this can be checked using the JISC detection service, Turnitin®.

- I understand that the work cannot be assessed unless both hard copy and electronic versions of the work are handed in.

- I declare that all material presented in the accompanying work is entirely my own work except where explicitly and individually indicated and that all sources used in its preparation and all quotations are clearly cited.

Should this statement prove to be untrue, I recognise the right of the Board of Examiners to recommend what action should be taken in line with UCL's regulations.

Signed   Sumin Chung

Print Name   Sumin Chung

Dated   20 / 08 / 2015

Title	Date Received	Examiner	Examiner's signature	MARK

Measuring Human Chromosomes using Confocal microscopy and Ptychography

Sumin Chung

Supervised by Prof. Ian Robinson

Dr. Fucai Zhang

Dr. Graeme Morrison

Submitted for the requirements for the degree of MSc at UCL

University College London

Department of Physics and Astronomy

August 2015

Abstract

The images of human metaphase chromosomes were obtained by using confocal microscopy and ptychography. The volumes and of each chromosome were measured from the images obtained by each method and the chromosomes were sorted into groups by size. Confocal microscopy is widely being used in the field of biology, physics and material science together with electron microscopy. The volume of chromosome was measured using height difference from confocal microscopy and height integration method from an imaging software. Ptychography is a novel type of lensless diffraction imaging method which reconstructs the image of an object from overlapping diffraction patterns made by multiple scans, which enables to overcome the limitation of CDI. From ptychography, the volume of the chromosome was calculated using the phase shift and refractive index. In this study, the volume of chromosome was measured by confocal microscopy to be compared and analysed with the results from ptychography. The volumes were compared with the contained number of basepairs for each chromosome.

Ⅰ. Introduction

Study of human chromosomes is focused on unraveling the unknown structure of DNA double strand by researching the fine structure of chromosome. This can be achieved by observing the chromosome with high resolution imaging techniques. Imaging has done important role in finding known aspects about the fine structure in biology. Xray diffraction patterns allowed the discovery of the DNA double strand, visible light and fluorescence microscopy allowed diagnosis of diseases by observing the structural change of genome position. [1]

The diameter of a nucleosome is about 10nm. H1, the 5th histone acts to form the fibre with diameter of 30nm, wrapped in a helical form with 6 nucleosomes per rotation. The precise molecular structure of this 30nm chromatin fibre is controversial. Attempts have been made to construct structural models for this, but the resolution for the scale of 30nm is not being achieved by most of the 3D imaging methods. The answer for this model is not

2

concluded yet and the arrangement of DNA and histones inside the fibre is still debated.[2]

Unlike transmission electron microscopy(TEM) whose penetration depth is smaller than the size of the chromosome, scanning electron microscopy(SEM), one of the most commonly used method in observing fine structures, can observe the chromosome fully but can only obtain two dimensional information since it senses only the surface, not the inner structure. [3]

Lensless imaging is a high resolution microscopy and is applied to such fields that face difficulties with optical imaging. The benefit of not using lens is that the resolution of the image depends only on the diffraction and dose limits. When coherent beam is scattered to an object, the resulted diffraction patterns are recorded. The recorded diffraction pattern reconstructs an image through certain algorithms that iterate between real and reciprocal space. This reconstruction is possible via field propagation method, only when both the intensity and the phase of the diffraction pattern are known. As one can only obtain information about the intensity from typical diffraction, the main interest of lensless imaging is recovering of the phase information from recorded diffraction patterns.

CDI allows the observation of the inner structure of the object and provides information about the structure of three dimensional object. Compared to conventional imaging, the advantage of lensless imaging is that it does not require high quality lens that satisfies atomic scale wavelengths. Therefore it is used to provide the image with good quality in imaging techniques that use illuminations with very short wavelength such as ultraviolet or Xray. It is also used to optimize the size and to reduce the cost for the device. [4]

Ptychography reconstructs the image from coherent diffraction using coherent illumination like CDI. While CDI analyses the intensity of the pattern from single diffraction, ptychography uses overlapping diffraction patterns produced by multiple scanning of the object. The advantage of ptychography is that it can reconstruct the image without the limitation of the sample size from overlapping diffraction patterns via specially developed algorithms whereas the size of the sample should not exceed the probe in CDI.

The goal of this study is to karyotype human chromosomes using quantitative information which is volume that is obtained by imaging methods, and to verify the potential of these imaging methods for fine structure studies in biology. Karyotype analysis is a technique used for sorting and identifying the chromosomes. Traditionally, chromosomes are differentiated by size and by a process called banding. Cells collected from an individual are arrested at the

metaphase stage of mitosis and then are stained with certain dyes that produce light and dark banding patterns. These bands reflect regional differences along the chromosome lengths in the amounts of the base pairs A and T versus G and C. the banding pattern for each chromosome is unique and consistent, which allows the identification of each of the 24 chromosomes. Human autosomal chromosomes are divided into seven groups (A-G) based on size and position of the centromere. [5]

In chapter 2, the important theoretical background of ptychography, coherence and the phase problem, are introduced. In chapter 3, the experimental procedures for confocal microscopy and ptychography with volume measuring process are explained. In chapter 4, the results for each experiment are compared and analysed with other studies. The overall evaluation and suggestions for further studies are introduced in chapter 5.

Ⅱ. Theory

1. Coherence

Coherent diffraction imaging(CDI) is a lensless imaging technique that reconstructs image from recorded diffraction pattern by using incident illumination with laser or Xrays. The beam that is used to make diffraction pattern should have coherence. Coherence is a property of waves when two wave sources have a constant phase difference and the same frequency. Coherence is divided into temporal coherence and spatial coherence. Temporal coherence, also called as longitudinal coherence, refers to a distance that phases of two waves that have different wavelengths travel from one point another point where they meet. Temporal coherence can be schematized as below.
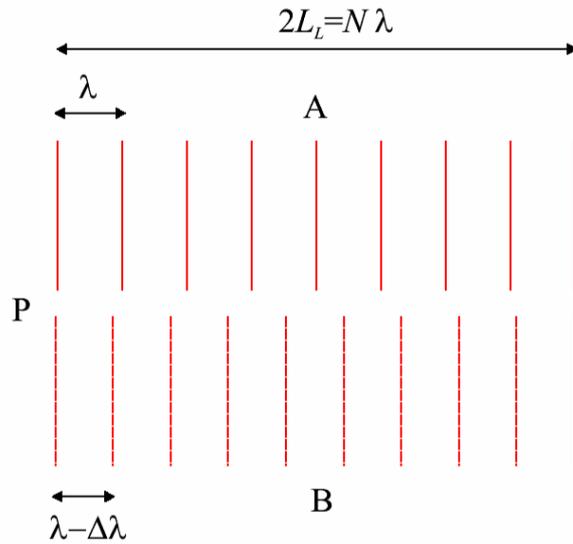
Fig1. *Schematic of temporal(longitudinal) coherence(L_L) of two waves A and B, each with wavelength of λ and $\lambda-\Delta\lambda$ respectively. [6]*

The two waves, A with wavelength λ and B with wavelength $\lambda-\Delta\lambda$, start from point P and travel to the right and meet at the second point. From the diagram above, the temporal coherence length can be expressed as $L_L = \frac{\lambda^2}{2\Delta\lambda}$. [1]

Spatial coherence is also known as transverse(or lateral) coherence and is determined by the paths of two waves emitted from the extremities of an extended source.
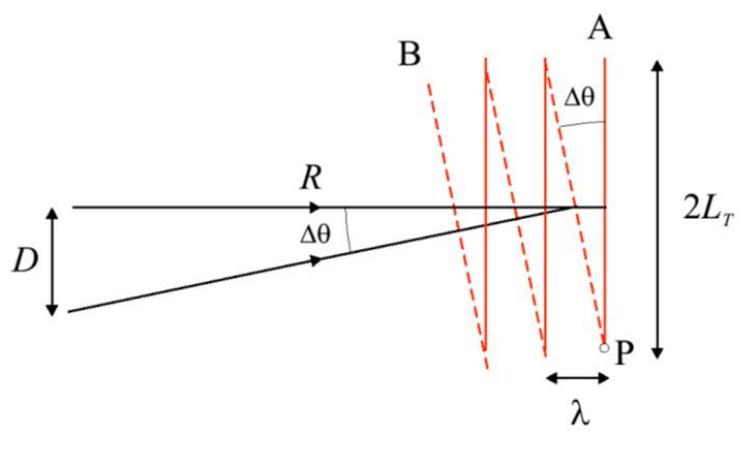

Fig2. *Schematic of spatial(transverse) coherence(L_T) of two waves A and B. [6]*

The phases of the two waves that started from each end of the source with size D meet at point P, which is aparted from the source by distance R. The distance from P to the point

where the two waves completely escape from each other in phase is twice the spatial coherence length. Therefore, the spatial coherence length is expressed as $L_T = \frac{\lambda R}{2D}$. [1]

2. The phase problem

In conventional imaging that uses lens to reconstruct the image, lens did the role of Fourier transformation without losing the phase information. In lensless diffraction imaging, the amplitude of the diffracted wave can be determined by measured intensity that is density of photons or electrons from the recorded diffraction pattern. However, since one can only obtain the modulus of the amplitude, the phase information is lost. The introduction of the iterative phase retrieval is one of the methods that are developed to retrieve the lost information about the phase in order to reconstruct the real space image. This uses multiple iteration of phase retrieval algorithm to recover the complex reconstruction of the sample. The modulus of Fourier transformation of the object represents the measured diffraction pattern and this condition forms the initial constraint set. Additional support constraint requires the reconstructed image to be zero outside the support. [7] Several algorithms to retrieve the phase from multiple diffraction patterns were developed. This approach has the advantage that it doesn't require finite number of object support. [8] The algorithm that uses these data is robust and shows faster convergence compared to when single diffraction pattern is used.
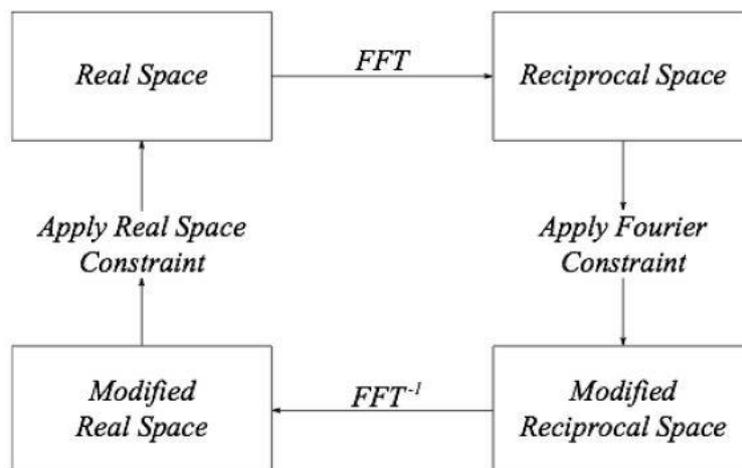
```
  ┌──────────────┐        FFT        ┌──────────────────┐
  │  Real Space  │ ────────────────► │ Reciprocal Space │
  └──────────────┘                   └──────────────────┘
         ▲                                    │
  Apply Real Space                    Apply Fourier
     Constraint                         Constraint
         │                                    ▼
  ┌──────────────┐       FFT⁻¹       ┌──────────────────┐
  │   Modified   │ ◄──────────────── │     Modified     │
  │  Real Space  │                   │ Reciprocal Space │
  └──────────────┘                   └──────────────────┘
```

The Fourier transform of the density distribution of the object *f(x)* is indicated as:

$$\mathcal{F}[f(x)] = \int f(x)\, e^{i2\pi(ux)} dx = F(u),$$

where *F(u)* is the diffracted wave, *x* is the real space coordinate and *u* is the reciprocal space coordinate. Similarly, the inverse Fourier transform is given as:

$$\mathcal{F}[f(u)] = \int f(u)\, e^{i2\pi(ux)} du = F(x)$$

With the help of Fast Fourier transform(FFT), we can solve discrete Fourier transform as follows:

$$F(u) = \sum_{x=0}^{N-1} f(x)\, exp(-i2\pi u * \frac{x}{N})$$

$$f(x) = N^{-2} \sum_{u=0}^{N-1} F(u) exp(i2\pi u * \frac{x}{n})$$

Then the diffracted wave *F(u)* can be written in discrete form:

$$F(u) = |F(u)| exp[i\psi(u)],$$

Where ψ represents the phase.

By conducting the inverse Fourier transformation of the intensity which is given from the square of $|F(u)|$, we can get the sample image function, *f(x)*, and thus find the phase information. [9]

Rodenberg and Faulkner (2004) proposed Ptychographical Iterative Engine(PIE) as a way that enables the measurement of spatially extended sample and verified the functional stability of the method. [10] The weakness of PIE is that it requires accurate model function. This spends very long time and at worst it could be impossible to measure the correct phase information of the probe function. Maiden et al. (2009) proposed an improved version of PIE algorithm, that is called extended Ptychographical Iterative Engine(ePIE), to resolve this problem.

2.1.   ePIE

ePIE is one of the commonly used algorithms in ptychography. It is extended to update both

the probe and the object functions during reconstruction in order to overcome the experimental difficulty that earlier algorithms for ptychography which use require very accurate known probe for successful retrieval had. The update of both function helps successful and fast convergence of the algorithm.

Dataset of ptychography is obtained from J far field diffraction patterns. During the measurement of each diffraction pattern, the probe and the sample move relative to each other in order to illuminate each region of the sample. The measurement keeps going until the area of interest of the sample is covered by several number of overlapped probe positions. The overlap required to get a good result is about 60~70 percent.

The interaction between the sample and the probe can be expressed by multiplication of both functions. The wavefront passed through the sample is called exit wave and it can be expressed by Fourier transformation when the wavefront is transmitted to the detector plane. The following flowchart represents the process of updating the probe and the object functions from $P_j(r)$ and $O_j(r)$ to $P_{j+1}(r)$ and $O_{j+1}(r)$.
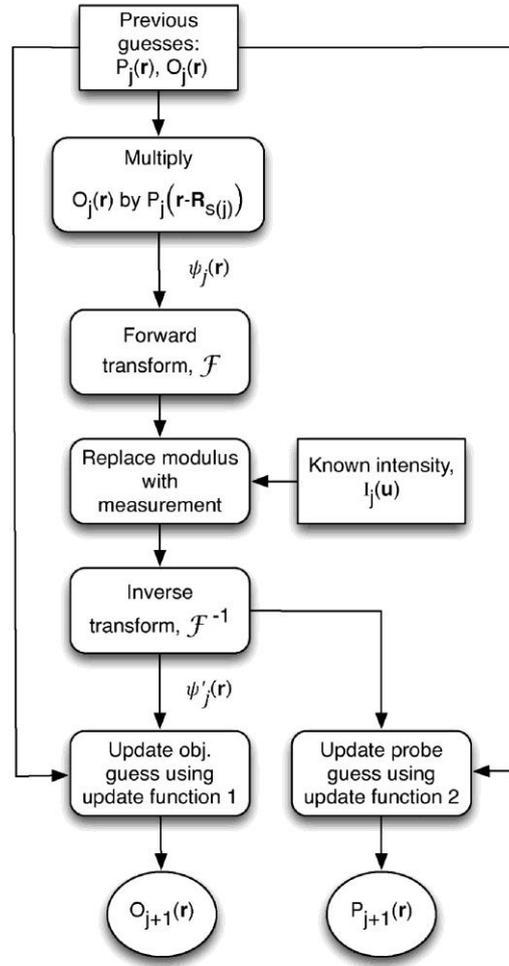
Fig4. *Flowchart of ePIE algorithm. At j=0, initial guess for the sample is provided as the free space guess and for the probe is provided as a support of roughly the correct size for the probe. [7]*

The initial guesses of the object and the probe function, i.e. $O_0(r)$ and $P_0(r)$ are necessary in order to start the algorithm. In general, $O_0(r)$ is determined as the free space guess of the sample and $P_0(r)$ is determined as rough size of illuminating region of the probe.

The exit wave guess is obtained by multiplication of current object guess and shifted probe guess.

$$\psi_j(r) = O_j(r)P_j(r - R_{s(j)})$$

The modulus of Fourier transform of the exit wavelength is replaced by the modulus of the s(j)th diffraction pattern as follows:

$$\psi_j(u) = \sqrt{I_{s(j)}(u)} \frac{\mathcal{F}[\psi_j(r)]}{|\mathcal{F}[\psi_j(r)]|}$$

As a result, the exit wave is expressed in the form of inverse Fourier transformation as follows:

$$\psi_j'(r) = \mathcal{F}^{-1}[\psi_j(u)]$$

In this way, each of the diffraction patterns is used to update both the object and the probe function for each step, which allows fast convergence.

From this result, the object and the probe guess are updated using the following two update functions. The update function 1 is as follows:

$$O_{j+1}(r) = O_j(r) + \alpha \frac{P_j^*(r - R_{s(j)})}{\left|P_j(r - R_{s(j)})\right|_{max}^2}(\psi_j'(r) - \psi_j(r)),$$

where α is a coefficient that adjusts the step size of the update. The update function 2, which is for updating the probe can be expressed in the same way. Here β is used to adjust the step size of the probe update.

$$P_{j+1}(r) = P_j(r) + \beta \frac{O_j^*(r + R_{s(j)})}{\left|O_j(r + R_{s(j)})\right|_{max}^2}(\psi_j'(r) - \psi_j(r))$$

Single iteration ends when all of the diffraction patterns from $s(1)$ to $s(j)$ are used to update the object and the probe. [7]

2.2.   pcPIE

In ePIE, the resolution of the reconstruction is determined by how accurate the movement of the probe relative to the object can be performed experimentally. Maiden et al. (2012) proposed position-correcting Ptychographical Iterative Engine(pcPIE), an extended version of ePIE, as a way of correcting this experimental error.

pcPIE iteratively updates both the object and the probe functions, but it additionally corrects the measured position of the specimen by updating the position correcting vector. The following flowchart shows how single iteration of pcPIE works.
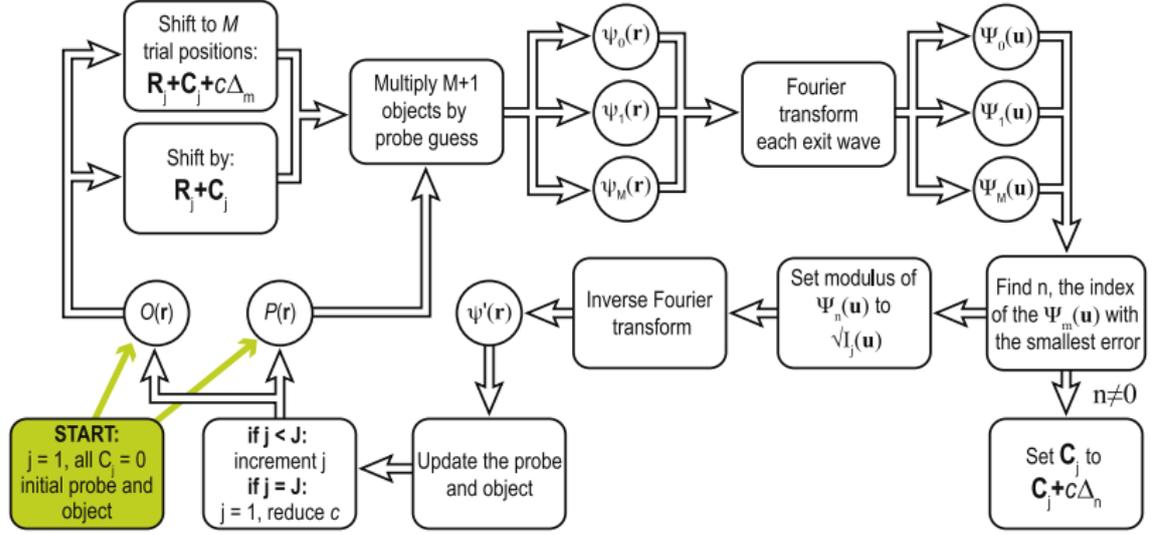
Fig5. *Flowchart of pcPIE algorithm. [11]*

The estimated initial wavefront ψ_0 can be expressed as follows:

$$\psi_0(r) = O\left(r - \left(R_j + C_j\right)\right) P(r),$$

where R_j is the sample position for each diffraction pattern that is measured inaccurately. C_j is the corrected position for each, which starts from zero and is updated by pcPIE algorithm.

The estimated wavefront for m=1, …, M is given as follows:

$$\psi_m(r) = O\left(r - \left(R_j + C_j + c\Delta_m\right)\right) P(r),$$

where $c\Delta_m$ indicates the trial offset for each. The Fourier transformation of this to result in as follows:

$$\Psi_m(u) = \mathcal{F}[\psi_m], m = 0, \dots, M$$

The error is obtained by comparing the modulus of $\Psi_m(u)$ with $\sqrt{I_j(u)}$:

$$E(m) = \sum_u \left| |\Psi_m(u)| - \sqrt{I_j(u)} \right|^2$$

The modulus of $\Psi_m(u)$ that represents the smallest error value is replaced by $\sqrt{I_j(u)}$, which is followed by inverse Fourier transformation to produce updated estimate of the wavefront.

$$\psi'(r) = \mathcal{F}^{-1}\left[\sqrt{I_j(u)} \frac{\Psi_n(u)}{|\Psi_n(u)|}\right]$$

Current correction vector C_j is updated by $c\Delta_n$ when smaller error is made.

$$C_j \leftarrow C_j + c\Delta_n$$

The object and the probe functions are updated using the updated wavefront and the correction vector.

$$O\left(r - \left(R_j + C_j\right)\right) \leftarrow O\left(r - \left(R_j + C_j\right)\right) + \alpha \frac{P^*(r)}{|P(r)|^2_{max}} (\psi'(r) - \psi_n(r))$$

$$P(r) \leftarrow P(r) + \beta \frac{O^*(r - \left(R_j + C_j\right))}{\left|O(r - \left(R_j + C_j\right))\right|^2_{max}} (\psi'(r) - \psi_n(r))$$

Single iteration of the algorithm ends when this process is applied for each of j patterns.

pcPIE algorithm is verified to improve the resolution and the quality of the image by modifying the position errors generated during the data collection process. [11] In the ptychography experiment performed in this study, ePIE and position correction were taken to be used as algorithms.

Ⅲ. Experiment

1. Confocal microscopy

1.1. Slide preparation

Four chromosome spreads, two of which are stained with Platinum Blue(PtBlue) and two of which are unstained, were taken to be measured from each slide. Each spread are human metaphase chromosomes and are in the different stage of condensation. PtBlue was used as dye to increase the mass and make the scattering better since chromosomes don't produce enough signal in Xray or SEM experiments. Therefore two kinds of slides were prepared to see if the stain changes the volume of the chromosome by any reason such as damaging the

chromosome structure in laser experiments.

The chromosomes were obtained from Yoruba cell line. Each spread contains 22 pairs of autosomal chromosomes and X, Y of sex chromosomes. 30μl of the sample liquid was dropped from height and 30μl of 10mM PtBlue was added for staining. The slide was covered by a cover slip and then placed for 30 minutes in dark. After then the cover slip was removed from the slide, which was followed by washing with distilled water for 5, 10, 15 minutes in sequence. Another slide was prepared by simply dropping 20μl of the sample liquid from height. The prepared slides were checked from fluorescence microscope to see if the chromosome spreads are well observed.

1.2. Confocal microscopy

Once the spreads are checked to be well observed by fluorescence microscope, they are observed again under confocal microscope to find the positions of good spreads. Good spreads for human karyotype analysis that selected were the ones that all of the 46 metaphase chromosomes are well spreaded, away from one another with decent distance to be easily measured by both confocal microscopy and ptychography. Position mapping of the chromosome spreads was needed since the karyotyping performed via observation from confocal microscopy and the volume measuring will be compared and analysed in ptychography with the same spread. Therefore a slide with same size and shape attached with thin and transparent grid map was prepared. Using the LEXT software program provided by Olympus OLS4100 confocal laser microscope, spreads can be searched with magnification of 2.5x to 100x by mapping. The grid coordinates of the good spreads were checked and recorded.

LEXT can perform volume measurement itself using height values of manually selected region. In this way the result can come out with great error due to the instability of the laser setup and lack of standard for setting the correct threshold between the background region and the chromosome. The laser setup of confocal microscope is so sensitive that it can be affected by very small movement. Also the threshold of the area of interest(AOI) is set by human and the difference of one pixel selection can result in huge difference in volume. Therefore there can be difficulty with inaccuracy to perform karyotyping based on this program only.

1.3. ImageJ software

The height map of each pixel of the spread checked from confocal microscope with resolution of X100 was saved as spreadsheet text image (.csv) from LEXT and was measured in ImageJ program. ImageJ converts the different heights of each pixels into brightness contrast and allows to calculate the volume per each selected region. Followings represent the process of obtaining the volume using confocal microscope and ImageJ.
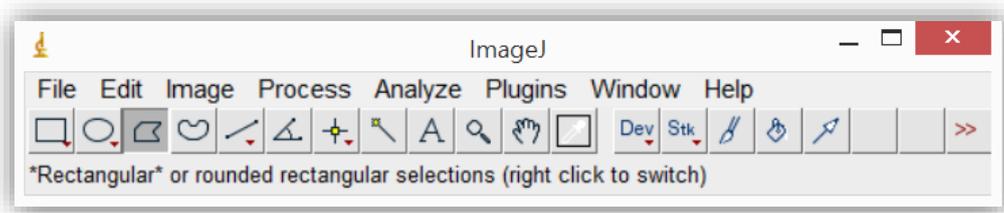
Fig6. *The height for each pixel saved from LEXT can be imported as a text image in ImageJ to get the volume from the image.*
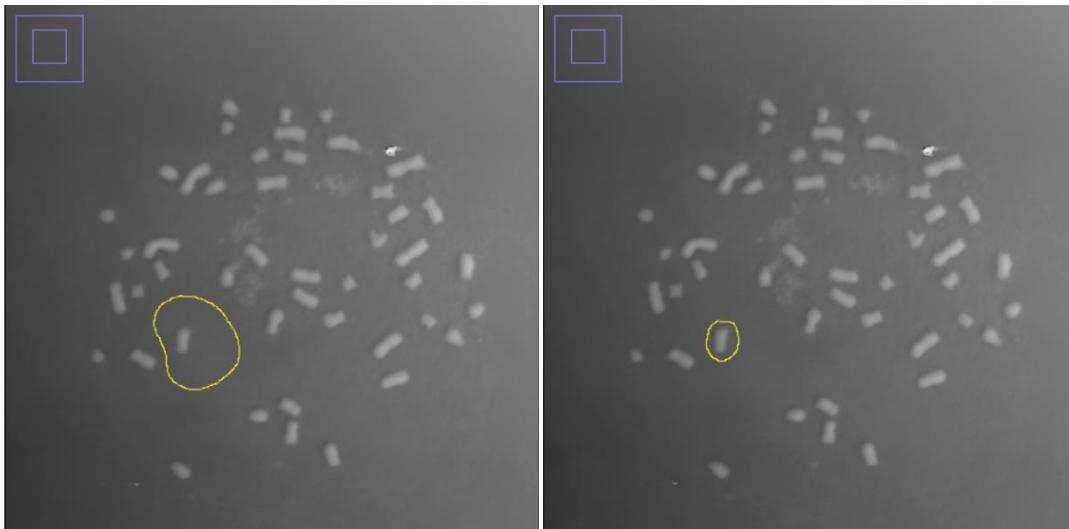
Fig7. *Widely selected region of interest, narrowly selected region of interest.*
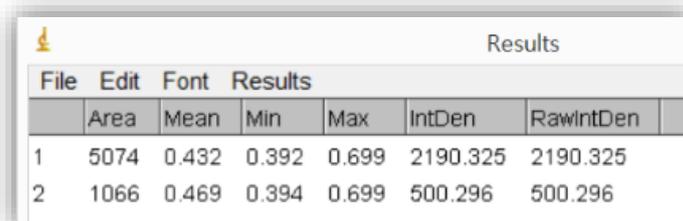
	Area	Mean	Min	Max	IntDen	RawIntDen
1	5074	0.432	0.392	0.699	2190.325	2190.325
2	1066	0.469	0.394	0.699	500.296	500.296

Fig8. The first row, 1, indicates the measurement for region of upper left, the second row, 2, is for upper right in Fig7.

Area of selected region, mean height, minimum and maximum height, density that is volume can be known from ImageJ measurement as can be seen from above. The volume is obtained by mean height multiplied by area.

Area	Height(μm)			Volume	
(pixel*pixel)	Mean	Min	Max	IntDen	RawIntDen
				=(area*mean height)	
5074	0.432	0.392	0.699	2190.325	2190.325
1066	0.469	0.394	0.699	500.296	500.296
Area difference				Volume difference	
=5074-1066				=2190.325-500.296	
4008				1690.029	
Mean background height		Background volume		Chromosome volume	
		(pixel*pixel*μm)		(pixel*pixel*μm)	
=1690.029/4008		=0.421663922*5074		=2190.325-2139.522741	
	0.421663922	2139.522741		50.80225898	
		Chromosome volume(μm^3)			
		=50.80225898*0.125*0.125			
		0.7937853			

Fig9. Process of calculating the chromosome volume in Excel.

To select the exact area where the chromosome is placed, the method of subtracting background from selected region was used for decide the volume of chromosome. First, select a wide nearby area(1) including certain chromosome and get the measurement. Then select smaller nearby area(2) and repeat the same job. From these two data of measurements, the mean height of the background where the certain chromosome is located can be obtained from the area difference and volume difference. The background volume is decided by multiplying the area of wide nearby region and the mean height of the background. Finally, the volume of the chromosome can be obtained by subtracting this from volume of (1).

From ImageJ, the measurement of area is shown in pixel unit whereas the height

values are presented in micrometer unit.

Fig10. Height map(.csv) obtained from confocal microscope software. X and Y resolutions are 0.125μm. Red box below from the data line shows the height(in micrometer) for each pixel(each cell in the spreadsheet).

Therefore, the pixel measurement for selected area should by multiplied by 0.125^2 to be represented in micrometer unit.
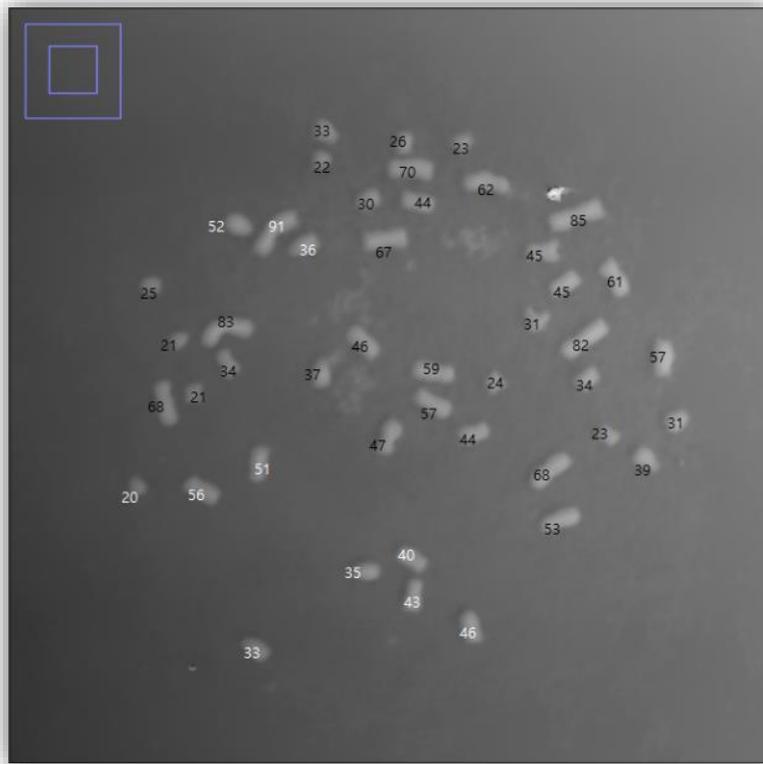
Fig11. *The volumes of one chromosome spread obtained by ImageJ method. Here the volumes are represented in pixel(x)*pixel(y)*μm(z, height) unit.*

This method doesn't require the exact threshold setting of the background and the chromosome itself to measure the volume, which reduces the difficulty from LEXT confocal microscope software.
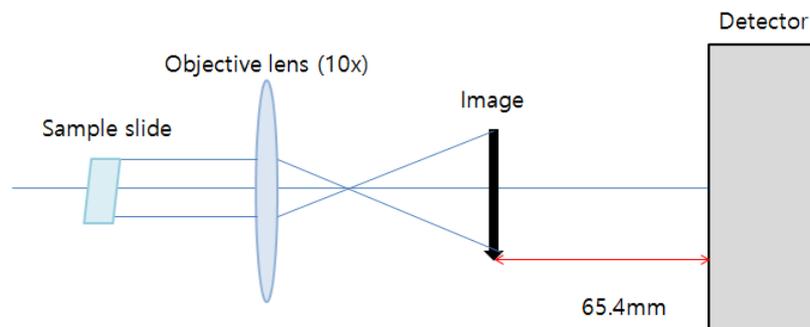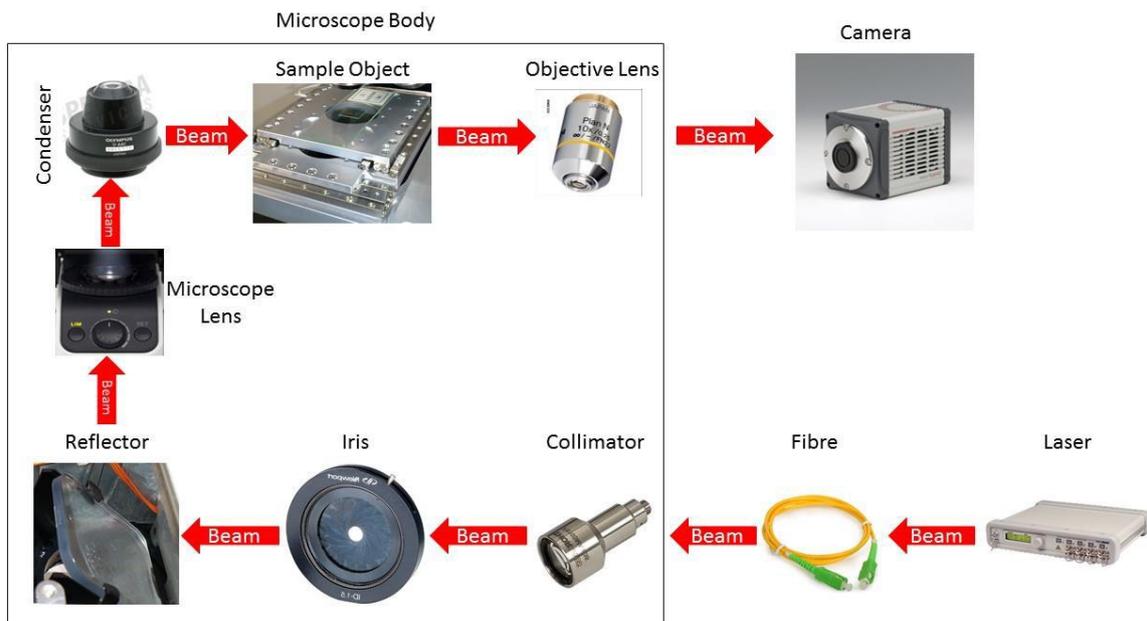
2. Ptychography

2.1.   Image reconstruction

The setup used in this experiment is shown below.
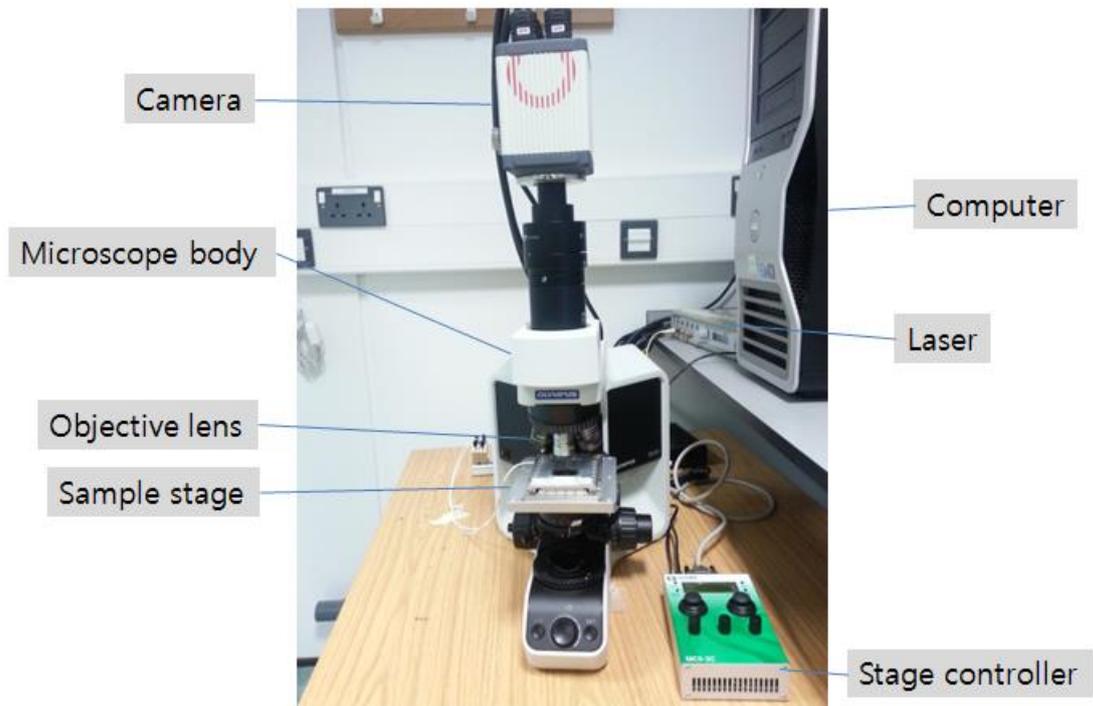
Fig12. *Ptychography setup*

Fig13. *Diagram of part of the setup (from the slide to the detector)*

Fig14. *Detailed demonstration of the hardware [12]*

Wavelength of 638nm, power of 15mW laser was emitted from the laser set. It passed through the condensing lens and the pinhole to be concentrated. A pinhole with 0.16mm diameter was used which determined the size of the illumination. The magnification of the objective lens used was 10x. The sample stage was placed under the objective lens and the position of the sample was adjusted by SmarAct MCS-3C stage controller. The stage controller can move the stage toward x and y direction to search the coordinate position.

First, the sample was checked from HCImageLive software on the connected computer to see if the camera is operating well. After then data was recorded using specially developed command set from Matlab software.

From confocal microscopy, thin grid was attached on the other side of the slide to find the position of the chromosome spreads. This caused disturbance since it was assumed that there is a unique focal plane on the sample in this reconstruction. Therefore another slide with the exact same position of the grid attached was prepared to be used to find the coordinate by moving the sample stage and then to be switched by the chromosome sample slide. Thibault (2007) explains this approximation that assumes the sample as a two dimensional object to obtain a good result. [8] If the sample isn't a two dimensional object, several focal planes would be produced and the average value of the planes will be assumed which can decrease the resolution of the reconstruction. [13]

19

During the process of making the slide for the grid, unstained sample slide was destroyed while stacking two slides to mark the exactly same position for the grid. When one sample is used for several different experiments, slide should be treated especially with care. Also, for the quality of the reconstruction, it is important to be careful not to contaminate the sample slide with the dust.

The chromosome spread is very small and weak as a sample in ptychography, and it is difficult to get a good quality of reconstruction with the probe for strong samples. Therefore, step-by-step approach for finding a strong probe was needed to be used for a weak sample. The weak sample, i.e. chromosome spread, was imaged by using accurate probe that was reconstructed from the imaging of a strong sample. First, a part of an insect was used as a strong sample. Here the initial probe guess is determined by the pinhole. Then, the updated probe from this step was used with the diffuser to image the strong sample. Finally, the strong probe obtained from the former step was used to image the weak sample. The role of the diffuser was to improve the resolution of the reconstruction by increasing the numerical aperture. [13]

The diffraction patterns were recorded after the spread position was adjusted using the stage controller. Round shaped probe sized 0.16mm was used with step size of 0.02mm, and the scanning range was 0.2mm*0.2mm. Total 84 points were scanned per each algorithm iteration.
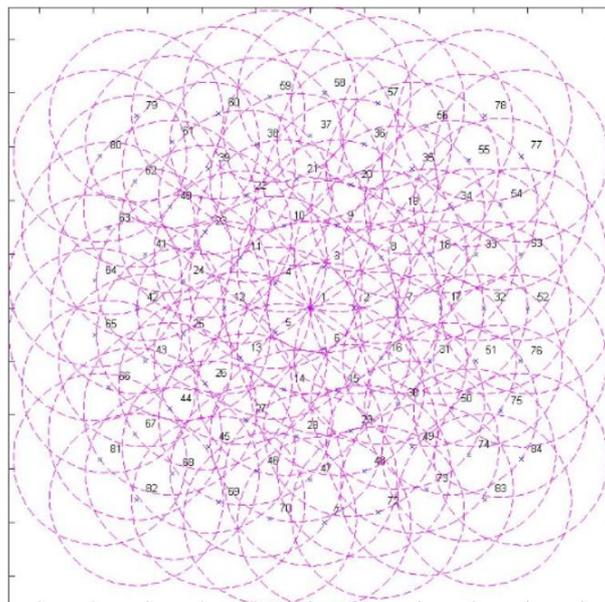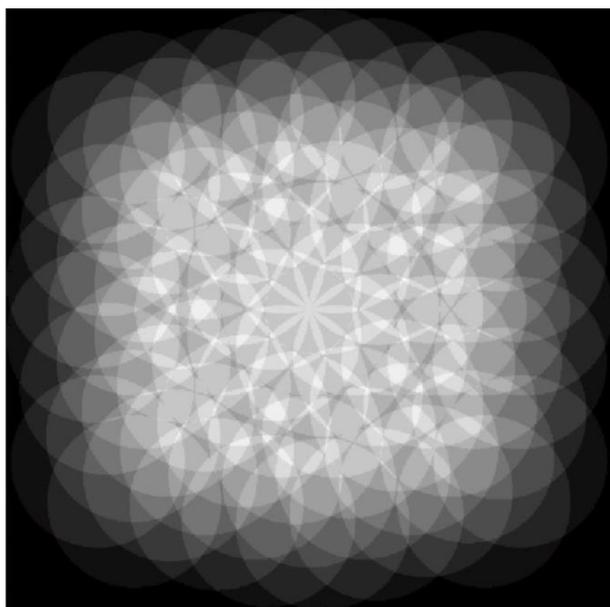
Fig15. *The position(upper) and the overlap map(lower) of the scan setting*

Finally, the recorded data was reconstructed via 60 iterations of algorithm. The probe was updated after 5 iterations. The followings are the images of four chromosome spreads reconstructed by ePIE and pcPIE. Following are the reconstructed images for S1 and S2 spreads.
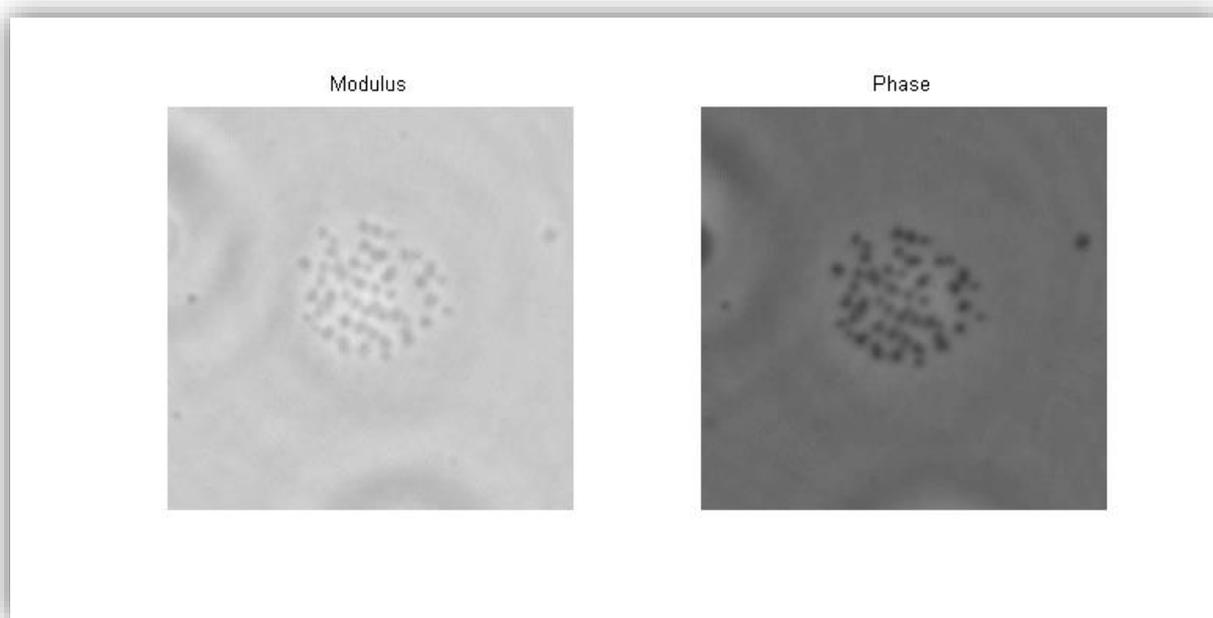
Fig16. *Modulus and phase images of S1(upper) and S2(lower)*

2.2. Volume measurement

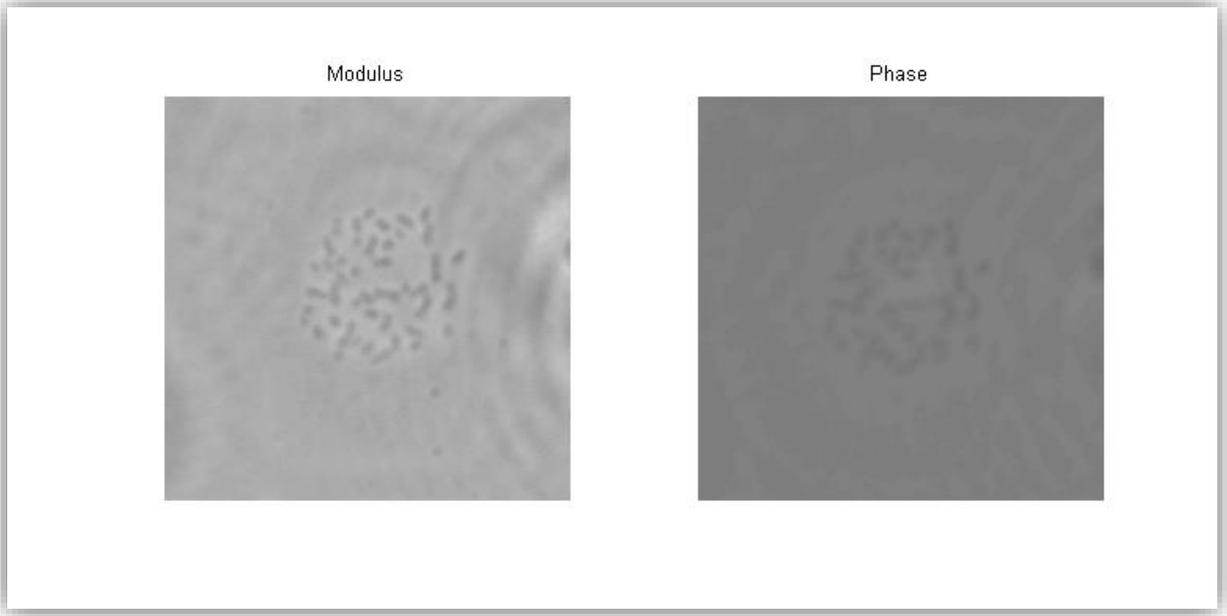The mass calculation used in the study of Laura et al. (2015) was applied for volume measurement. In the study, they obtained the chromosome mass and conducted partial karyotyping.

Since the phase images are discretised into pixels, each pixel can be expressed as a position vector *(x,y)*. The object function *O* is described as follows:

$$O = A(x, y)e^{-i\varphi(x,y)},$$

Where *A(x,y)* is the amplitude of the wave function, and the phase is given by $\varphi(x, y)$.

The phase through each point $\varphi(x, y)$ is related to the projected mass density *dm/dxdy* as follows:

$$\varphi(x, y) = \frac{2\pi(n - n_0)}{\lambda\rho} \frac{dm}{dxdy}$$

where *n* is the refractive index of the object, n_0 is the refractive index of the background, λ is the wavelength, and ρ is the density of the material. If the background is air then $n_0 = 1$. Hence the phase in this experiment can be expressed as:

$$\varphi(x, y) = \frac{2\pi\delta}{\lambda\rho} \frac{dm}{dxdy}$$

The mass of the chromosome is found by integrating the phase across the

chromosome area:

$$m = \int dm = \frac{\lambda \rho}{2\pi(n-n_0)} \iint \varphi(x,y)\, dx dy.$$

The integral above can be expressed in the form of addition of the phase for each pixel:

$$m = \frac{\lambda \rho}{2\pi(n-n_0)} \Sigma_n \varphi_n \, p,$$

As the phase image we get will be in the form of discrete pixels. φ_n represents the phase value of the image at pixel n and p is the pixel resolution. [14]

From this, we can obtain the volume to compare with the confocal microscopy results. Volume is determined by mass divided by density. Hence,

$$V = \frac{\lambda}{2\pi(n-n_0)} \Sigma_n \varphi_n \, p.$$

2.3.   ImageJ software

The phase image of the object was exported as text image (.csv) file from the Matlab, to be used in the same way as in confocal microscopy. The selected area, the phase information(average, minimum, maximum phase values) of the area of interest were able to be obtained in ImageJ.
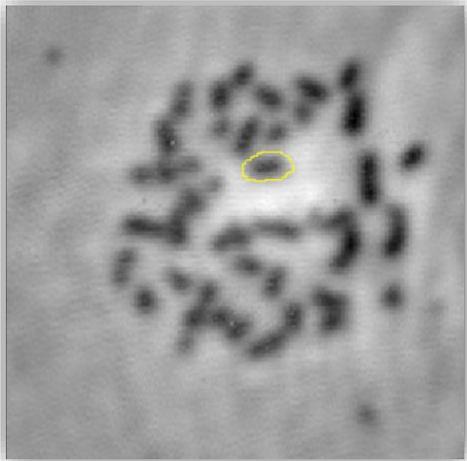
Fig17. *Imported phase image in ImageJ(upper) and the measurements of the selected chromosome area(lower).*

* The selected region is consisted of 295 pixels (pixel size: 313nm), mean phase value is shown as a negative number, as this is the chromosome region.

**The minimum phase value would be the highest(thickest) pixel inside the chromosome.

***From the maximum value which is shown as a positive number, it can be known that the region is including the background pixel.

The phase value has a negative number in chromosome area and a positive number in the background area, which makes it impossible to treat this in the same way it was done in the height map to get the volume. Therefore manual threshold setting was needed for each chromosome. It will cause errors in the volume measuring when the background region is selected in the threshold setting as in **Fig17.** To avoid this, all the pixels which are represented to have positive phase value are eliminated.
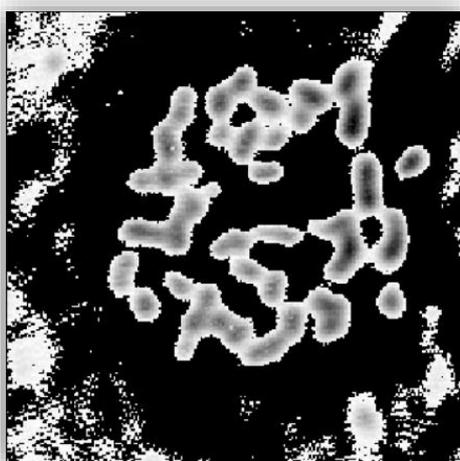
Fig18. *The darkened region in the background is treated to be "None", which means no pixels are considered. Therefore the pixels of the background, i.e. pixels with positive phase value, will not be included in the measurement which makes it easier to set the threshold.*

Also, since the background phase values will vary across the whole spread, the mean phase for the nearby region where each chromosome is place was measured individually and subtracted from the mean phase value of each chromosome. As a result, the mean phase value for only the chromosome was obtained to determine the thickness and this was used to calculate the volume by multiplication with the area.

As can be seen from the phase map image above, there is high probability of error when the chromosomes are closely located to one another. Since the phase reconstruction is usually not as sharp as the modulus image, the threshold setting for these cases was done refer to the modulus image.

Ⅳ. Results

1. Confocal microscopy

The measured volume of four sample spreads are shown from 1 to 46, lined up by size as below. Karyotyping is performed based on the size and shape of each chromosome, but here the shape wasn't considered in order to see whether it is possible to karyotype only by volume(i.e. size).

Number	Volume				Group
	Stained 1		Stained 2		
1	1.35	1.26	1.53	1.48	A
2	1.20	1.20	1.35	1.32	
3	1.06	1.06	1.17	1.17	
4	1.01	0.969	1.10	1.06	B
5	0.953	0.953	0.953	0.938	
6	0.938	0.938	0.891	0.859	C
7	0.922	0.891	0.828	0.828	
X	0.766		0.828		
8	0.766	0.766	0.813	0.781	
9	0.719	0.703	0.750	0.750	
10	0.703	0.703	0.719	0.719	
11	0.672	0.625	0.688	0.672	
12	0.625	0.547	0.609	0.594	
13	0.531	0.516	0.578	0.578	D
14	0.516	0.500	0.563	0.516	
15	0.484	0.484	0.484	0.484	
16	0.469	0.422	0.484	0.453	E
17	0.422	0.391	0.438	0.422	
18	0.375	0.359	0.406	0.391	
19	0.344	0.344	0.344	0.344	F
20	0.328	0.328	0.313	0.313	
Y	0.313		0.281		
21	0.297	0.281	0.266	0.250	G
22	0.266	0.266	0.250	0.234	
sum	30.5		31.8		

Number	Volume				Group
	Unstained 1		Unstained 2		
1	1.42	1.32	1.53	1.43	A
2	1.29	1.28	1.37	1.35	
3	1.09	1.06	1.31	1.17	
4	1.06	1.04	1.17	1.10	B
5	0.968	0.953	1.10	1.03	
6	0.921	0.890	1.00	1.00	C
7	0.890	0.875	0.984	0.953	
X	0.828		0.953		
8	0.812	0.796	0.937	0.906	
9	0.734	0.718	0.890	0.890	
10	0.718	0.703	0.890	0.875	
11	0.703	0.687	0.812	0.781	
12	0.687	0.671	0.781	0.734	
13	0.625	0.609	0.671	0.671	D
14	0.578	0.562	0.671	0.609	
15	0.546	0.531	0.609	0.593	
16	0.531	0.515	0.578	0.546	E
17	0.515	0.484	0.531	0.515	
18	0.484	0.468	0.484	0.484	
19	0.406	0.390	0.484	0.437	F
20	0.375	0.359	0.421	0.390	
Y	0.359		0.375		
21	0.343	0.328	0.359	0.359	G
22	0.328	0.312	0.328	0.250	
Sum	32.8		36.3		

Table1. Autosomal chromosomes of four spreads sorted into groups by volumes measured from confocal microscope and ImageJ. Sex chromosomes X, Y were roughly determined by volume ordering.

The images of chromosomes grouped by volume are shown below. Sex chromosomes X and Y are grouped as group C and G, as these are difficult to determine based on measured volumes.

26
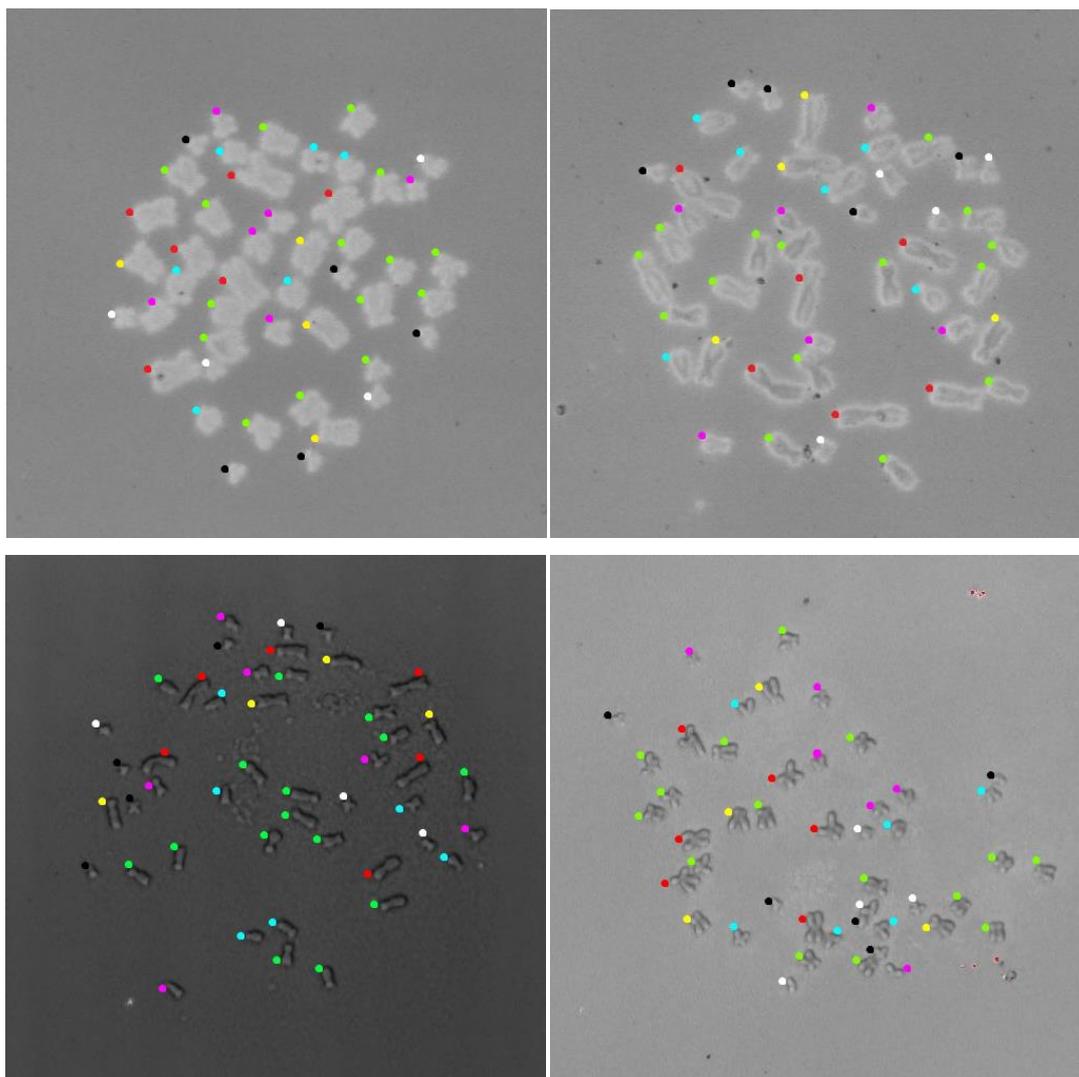
Fig19. *Chromosomes of four spreads (upper two stained with PtBlue, lower two unstained) grouped by volume ordering. The colour of each dot indicates the group in **table1**.*

Also, the lengths of chromosomes 1, 2 were measured from LEXT in order to see the relation between the condensation stage and measured volumes of each spread.

Spread	Length											
	S1			S2			U1			U2		
			avg			avg			avg			avg
Chromosome 1	6.76	6.13	6.45	8.25	7.40	7.82	4.52	4.50	4.51	3.55	3.45	3.50
Chromosome 2	5.92	6.05	5.98	7.44	8.43	7.93	4.46	4.22	4.34	3.85	3.58	3.71
Total volume (sum)	32.8			36.3			30.5			31.8		
Total volume (ImageJ)	49.6			70.0			77.0			46.1		

Table2. *Length of chromosome 1, 2 (the largest ones) for each spread(unit: μm). Total*

27

volume(sum) indicates the sum of the measured volume of each chromosome(unit: μm^3). Total volume(ImageJ) indicates the volume of the spread measured directly from ImageJ all at once(unit: μm^3).

From the table above, spreads with shorter measured length can be interpreted as less condensed ones. Therefore, comparing the lengths of stained spreads S1 and S2, S1 is considered to be more condensed. In the process of condensation and decondensation of the chromosome, it is assumed that the chromosome mass for each stage will be the same. That is, M_1=M_2 if M_1 is the mass for condensed chromosome and M_2 is for decondensed one. M_2 might be smaller than M_1, considering the protein-cutting by protease during decondensation. Also, considering the density for the condensed stage ρ_1 should be bigger than ρ_2, it is expected in theory that the volume for more condensed chromosome V_1 is bigger than V_2.

  Total volume(sum) and total volume(ImageJ) shows quite big difference but both indicates that S1 has smaller volume than S2. From unstained spreads U1 and U2, average length shows that U2 is more condensed, which means theoretically should have lesser volume, however U2 shows slightly larger number in total volume(sum), while much smaller number in total volume(ImageJ). There could have been some errors during the measurement considering there are not much difference in total volume(sum).

  Overall, it can be seen that S2>S1>U1>U2 in order by length, U1>S2>S1>U2 in order by total volume(sum), and S2>S1>U2>U1 in order by total volume(ImageJ). When we exclude U1 spread which shows huge difference between the two volume measuring methods, we can see the relationship between the length(that is, the condensation stage) and the volume, S2>S1>U2. Also, there are no noticeable difference observed between PtBlue stained and unstained ones considering the length of each spreads though the expected volume of stained spreads were bigger than unstained ones.

The volume measurement from confocal microscope was compared with known number of base pairs per each chromosome. The number of base pairs for 24 (22 homologous chromosomes + X, Y sex chromosomes) human chromosomes are known as follows:

Chromosome	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y	Total
Base pairs (10^6)	247	243	199	191	181	171	159	146	140	135	134	132	114	106	100	89	79	76	64	62	47	50	155	58	3079

Table3. *Estimated number of base pairs on each human chromosome. [15]*

From the measured volume for each chromosome, we can see how well it agrees the linear correlation between the genome size and volume of each chromosome.
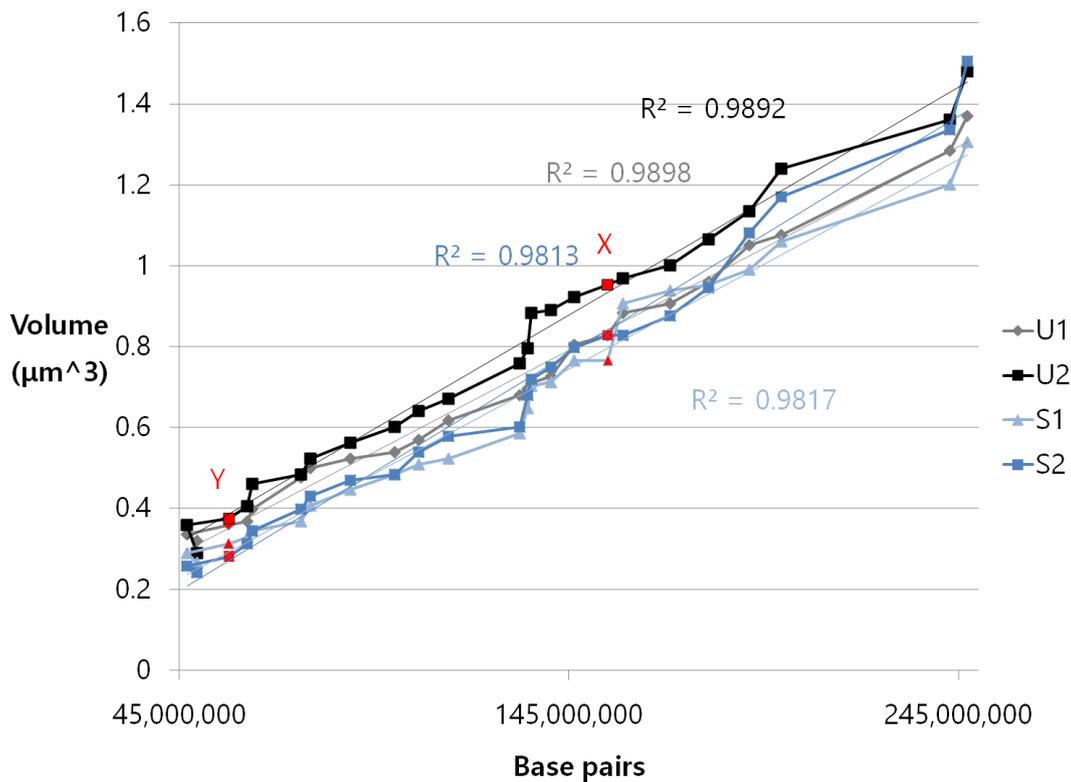
Graph1. *Correlation between the number of base pairs and volume of each chromosomes for the four spreads. The average R^2 value of spreads from U1 to S2 is 0.9855. Each of the 24 dots indicates the average volume of chromosome 1~22, the homologous chromosomes, except for the two red dots, which indicate X and Y chromosomes. X, Y are assumed to have the volume between chromosome number 7 and 8, 20 and 21, respectively.*

2. Ptychography

The following are karyotype results based on volume measurement from ptychography.

Number	Volume				Group
	Stained 1		Stained 2		
1	0.898	0.821	1.40	1.40	A
2	0.799	0.798	1.19	1.05	
3	0.771	0.752	0.945	0.937	
4	0.729	0.711	0.895	0.792	B
5	0.652	0.649	0.733	0.726	
6	0.646	0.577	0.688	0.649	C
7	0.526	0.515	0.647	0.615	
X	0.502		0.610		
8	0.493	0.483	0.574	0.571	
9	0.475	0.460	0.536	0.519	
10	0.458	0.434	0.514	0.511	
11	0.414	0.414	0.500	0.494	
12	0.401	0.387	0.493	0.483	
13	0.377	0.373	0.477	0.473	D
14	0.366	0.340	0.466	0.443	
15	0.335	0.324	0.436	0.425	
16	0.312	0.295	0.375	0.326	E
17	0.289	0.262	0.300	0.292	
18	0.258	0.235	0.277	0.277	
19	0.231	0.227	0.261	0.259	F
20	0.224	0.212	0.258	0.257	
Y	0.188		0.237		
21	0.152	0.151	0.209	0.175	G
22	0.145	0.0920	0.0980	0.0908	
sum	20.2		24.9		

Table4. Autosomal chromosomes of two spreads sorted into groups by volumes measured from ptychography and ImageJ. Sex chromosomes X, Y were roughly determined by volume ordering.

The volume of spreads with PtBlue stained only were able to be obtained since the slide for unstained two spreads were destroyed. Compared with the result from confocal microscopy, the volume measured from ptychography is smaller. Also it can be seen from the table 4 that the volume of chromosome in S2 is larger than S1.
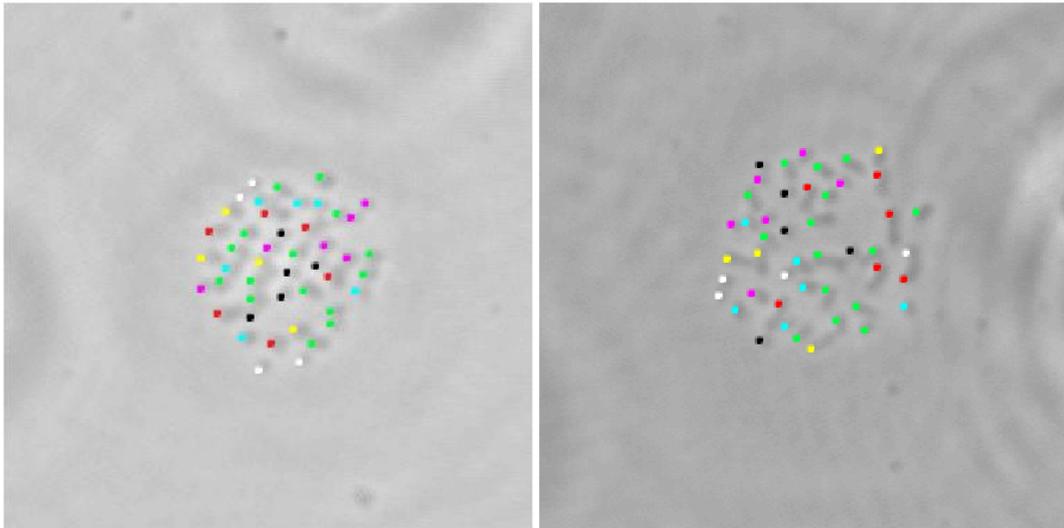
Fig20. *Chromosomes of two spreads (stained with PtBlue) grouped by volume ordering. The colour of each dot indicates the group in **Table4**.*

It can be seen that comparing Fig19 and Fig20, 22 and 23 numbers from the total 46 chromosomes were sorted into same groups in both measurements, for S1 and S2 respectively. There are 14 and 16 numbers of chromosomes showing correspondence in grouping from groups A~C, which indicate 25 chromosomes including chromosome number 1~12 and X chromosome. Also, 8 and 7 chromosomes are showing correspondence from groups D~G, 21 chromosomes including chromosome number 13~22 and Y. It can be known that the volume measuring becomes difficult as the size gets smaller.

In the same way as from confocal microscopy, the correlation between the number of basepair and volume is represented as follows.
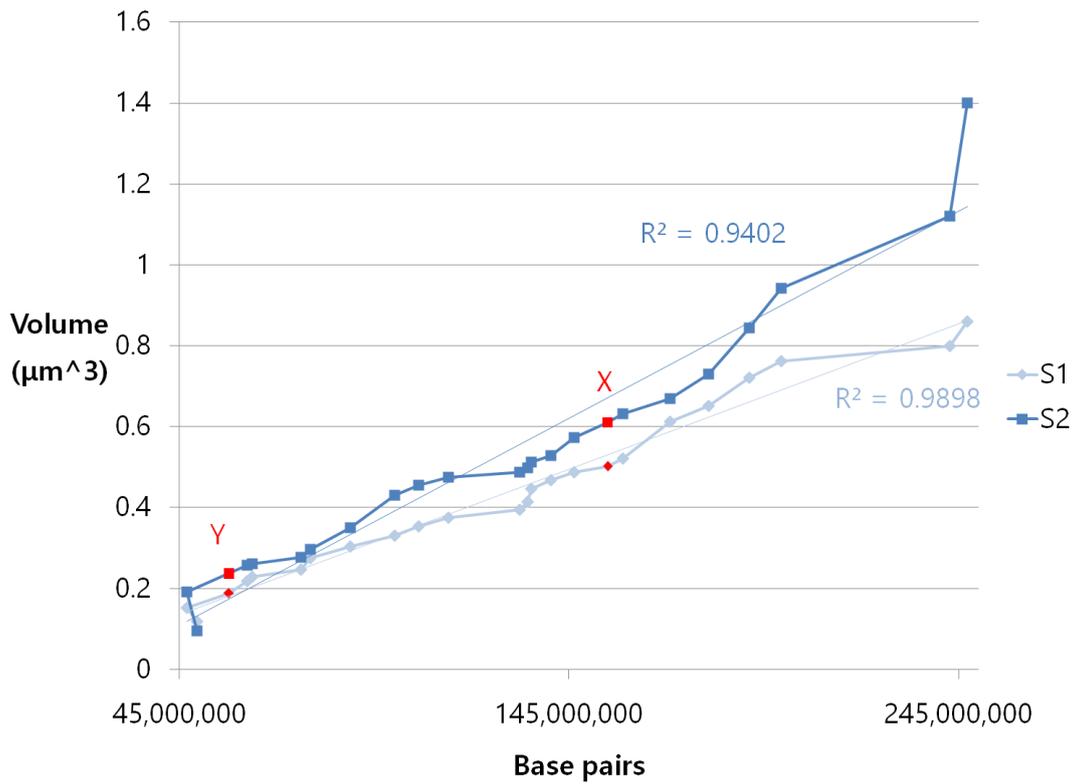
Graph2. *Correlation between the number of base pairs and volume of each chromosomes for the two spreads. The average R^2 value of spreads for S1 and S2 is 0.9650. Each of the 24 dots indicates the average volume of chromosome 1~22, the homologous chromosomes, except for the two red dots, which indicate X and Y chromosomes. X, Y are assumed to have the volume between chromosome number 7 and 8, 20 and 21, respectively.*

Ⅴ. Conclusion

The accuracy and the reliability of the results from confocal microscopy and ptychography will be able to verified afterwards by perfoming M-FISH. These quantitative information can be used complementarily in chromosome studies. Karyotyping can be conducted with accuracy from M-FISH, but by using different methods for measuring chromosomes we can obtain further knowledge for structural information about the chromosomes.

In this study, human chromosome spreads were imaged from confocal microscopy

and ptychography, and the expected numbers based on the volume measurements were given for each chromosome. The results from both experiments were used to check the correlation of them with the known number of base pairs. The supplementation to improve the resolution and to obtain further knowledge about the structure will be required.

The important thing I realised was how different it is when we perform an experiment with biological material compared to general physics experiments. Since the condensation stage for each chromosome spread is different, it was hard to directly compare the size and the volume with one another. This was also awared in the process of making mistakes while preparing biological samples and keeping them for several experiments. A year of time spent for learning scientific experiments from the beginning and studying biology and physics combined was fantastic. In Research Complex at Harwell I got to experience various kinds of optical instruments including microscopes and lasers, learn several practical methods for imaging and became interested in chromosome studies. It will be great to have opportunity to study further about 3D imaging in the future.

References

[1] L. Shemilt. 2014. Coherent diffraction imaging and ptychography of human metaphase chromosomes

[2] J. Langowski and H. Schiessel. http://www.researchgate.net/profile/Joerg_Langowski. Theory and computational modeling of the 30nm chromatin fiber.

[3] L. Shemilt, A. Estandarte, M. Yusuf, and I. Robinson. Scanning electron microscope studies of human metaphase chromosomes

[4] S. Witte, V. Tenner, D. Noom, and K. Eikema. 2014. Light: Science & Applications (2014) 3, e163; doi:10.1038/las.2014.44. Lensless diffractive imaging with ultra-broadband table-top sources: from infrared to extreme-ultraviolet wavelengths

[5] http://science.jrank.org/pages/3749/Karyotype-Karyotype-Analysis.html

[6] http://www.cmmp.ucl.ac.uk/~ikr/pub/talks/Robinson_Cardff_309.pdf

[7] A. Maiden, J. Rodenburg. 2009. Ultramicroscopy 109 (2009) 1256~1262. An improved ptychographical phase retrieval algorithm for diffractive imaging

[8] P. Thibault. 2007. Algorithmic methods in diffraction microscopy

[9] Q. Luo. 2013. Optical ptychographic phase tomography

[10] J. Rodenburg and H. Faulkner. 2004. Applied physics letters, volume 85, number 20. A phase retrieval algorithm for shifting illumination

[11] A. Maiden, M. Humphry, M. Sarahan, B. Kraus, J. Rodenburg. 2012. Ultramicroscopy 120 (2012) 64~72. An annealing algorithm to correct positioning errors in ptychography

[12] A. Azad. 2014. Optical ptychography imaging

[13] E. Verbanis. 2013. Study of chromosomes using ptychographic phase retrieval method

[14] L. Shemilt, E. Verbanis, J. Schwenke, A. Estandarte, G. Xiong, R. Harder, N. Parmar, M. Yusuf, F. Zhang, and I. Robinson. 2015. Biophysical Journal, Volume 108, February 2015, 706~713. Karyotyping human chromosomes by optical and x-ray ptychography methods

[15] http://vega.sanger.ac.uk/Homo_sapiens/Info/Index