Genet. Res., Camb. (1980), **36**, *pp.* 219–234
With 7 text-figures
Printed in Great Britain

Distribution of human chromosomes on the metaphase plate. Symmetrical arrangement in human male cells

By M. DOLORES COLL,* CARLOS M. CUADRAS† and JOSE EGOZCUE*

**Department of Cell Biology, Faculty of Science, and Instituto de Biología Fundamental, Universidad Autónoma de Barcelona, Barcelona, Spain*
† *Department of Genetics (Biostatistics), Faculty of Biology, University of Barcelona, Barcelona, Spain*

(*Received* 24 *August* 1979 *and in revised form* 18 *March* 1980)

SUMMARY

Two linear models have been devised and applied to the study of the distribution of human male chromosomes on the metaphase plate in preparations from lymphocyte cultures not treated with spindle poisons. Using these it has been found that the chromosomes are approximately distributed around a centre of symmetry and that the lines joining the centromeres of homologous chromosomes (i.e. segments) have the centre of symmetry approximately at their mid-point. Thus each chromosome mirrors the position of its homologue relative to the centre of symmetry. The position of each chromosome in the metaphase plate was found to be approximately constant relative to the centre of symmetry and the other chromosomes. The significance of these findings is discussed in relation to the hypothesis on the distribution of the chromosomes in interphase nuclei and to data on acrocentric associations, acrocentric–non-acrocentric associations and the frequency of the most common translocations.

(1) INTRODUCTION

The distribution and relative positions of metaphase chromosomes have been extensively studied. According to most published results, the distribution of the chromosomes on the metaphase plate seems to be non-random, and may be related to the somatic association of the homologues (Feldman, Mello-Sampayo & Sears, 1966; Egozcue, 1968; Fedak & Helgason, 1970 and Yoshida, 1974) or to some other type of ordered arrangement in interphase (Comings *et al.* 1968). Probably the most important differences existing among the studies conducted so far depend on the statistical treatment of the data and on the effect of spindle poisons.

Spindle poisons seem to be the most important agents in disturbing the normal relationships among chromosomes. In two different studies on the position of

0016–6273/80/2828–8080 $01.00

Muntjak chromosomes in metaphase published simultaneously, Cohen, Enis & Pfeifer (1972) using colcemid concluded that the distribution was random, while Heneen & Nichols (1972) who omitted its use observed the presence of somatic pairing. However, with some exceptions (Heneen & Nichols, 1972 and Juricek, 1975) this influence has not been taken into account in the studies carried out so far on chromosome distribution in metaphase (Barton & David, 1962; Miller *et al.* 1963*a*, *b*; Barton, David & Merrington, 1965; Galperin, 1968, 1969; Ockey, 1969; Hoo & Cramer, 1971; Cohen *et al.* 1972 and Sele *et al.* 1977).

The statistical methods used by most authors have been mainly based on the distance of Mahalanobis and in general have not taken into account the variability existing in the metaphase figures as a result of technical differences, projection of the chromosomes on the metaphase plate due to their spacial orientation, statistical errors and lack of a common point of reference for all metaphases studied.

To establish a statistical model that could be used in the study of the distribution of chromosomes in metaphase we established the following conditions: to avoid the possible influences of individual characteristics, the study would be carried out in 30 G-banded (Yunis *et al.* 1971) metaphases from lymphocyte cultures of 30 normal males, not treated with spindle poisons; well spread metaphases would be chosen at random under low-power; thus it was assumed that some of the figures would not contain a complete set of 46 chromosomes, because the number of metaphases is much lower than in regular preparations and some chromosomes would be unidentifiable due to superimpositions, or lost because of too much spreading; this possibility had to be taken into account in the model. In this paper we describe our results.

(2) MATERIALS AND METHODS

Lymphocyte cultures from 30 normal males were set up in a standard medium (Gibco 1-A). The cultures were not treated with spindle poisons. Preparations were made using a modification of the technique of Moorhead *et al.* (1960) after a 15 min hypotonic shock in 0·038 M-KCl. G bands were obtained using a modification of the method of Yunis *et al.* (1971).

To avoid the possible influence of individual characteristics on the final results, only one well spread, G-banded metaphase from each subject was used. The metaphases were selected under low-power to avoid a possible unconscious bias of the observer, and photographed. It was obvious that in well spread figures from cultures not treated with spindle poisons not all metaphases selected would contain a complete set of 46 chromosomes; however, the statistical model devised had taken into account this circumstance. When one or both members of a given pair were lost or unidentifiable this pair was not included in the calculations. Table 1 shows the number of metaphases and the chromosomes that were not taken into consideration or not found in each case.

In each metaphase, the coordinates x_i, y_i were established for the 46 centromeres

in relation to a pair of arbitrary axes. Each metaphase was then drawn using a computer plotter system programmed in Fortran IV level F. An example is given in Fig. 1.

The statistical model used can be summarized as follows:

Table 1. *Chromosomes present or not taken into consideration in each metaphase plate*

Chromosome no.	No. of metaphases: Present	No. of metaphases: Not taken into account
1	30	—
21	30	—
2	29	1
16	28	2
3, 4, 5, 7, 13, 14, 15	27	3
6	25	5
11, 12	24	6
12, 18, 19	23	7
17, 20	21	9
8	18	12
9	17	13
10, X, Y	15	15
Complete set	7	—

(i) *The problem*

Consider m photographs and 46 centromere positions corresponding to $q = 23$ chromosome pairs in each. We want to establish a model to objectively represent the mean position (mean of m photographs) of the chromosomes, with consequent margin of error.

The variability of the chromosomes depends on three factors: (A) The position of the chromosomes in a photograph is different for each of them. This is the systematic variability due to the different situation of each chromosome in the metaphase plate.

(B) The position of each chromosome varies from one photograph to another because it is impossible to take each picture according to a fixed reference, in a constant direction and with the necessary translation with respect to the origin.

(C) In any photograph, the position of the chromosomes shows a residual variability in relation to an ideal, fixed position. This may be due to technical problems, projection of the chromosomes on the metaphase plate, statistical errors, etc. and corresponds to the random variability of the experiment.

(ii) *Fixing the reference*

The main problem in elaborating a general model is due to the difficulty in identifying each chromosome individually and distinguishing it from its homologue. This difficulty can be avoided by taking for each pair (i,i') composed of

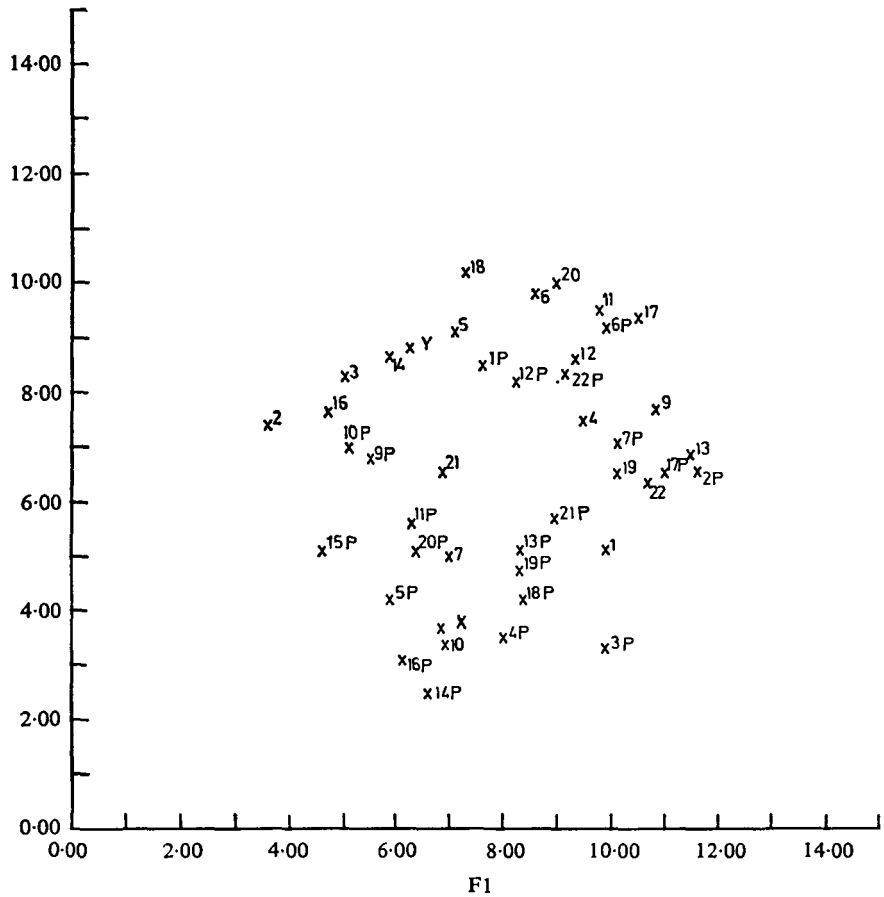


Fig. 1. Metaphase plate drawn using the PLOTTER CALCOMP-563 'of line' system programmed in FORTRAN IV level F.

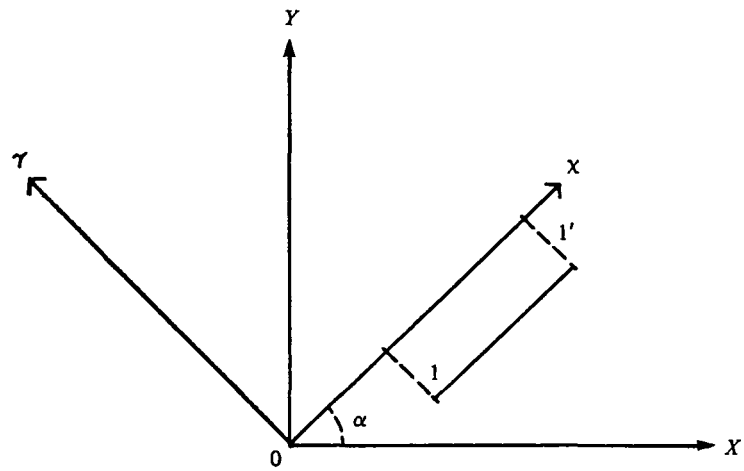
chromosome i and its homologue i' the following data: (a) mean points of each pair, and (b) segments (not oriented) joining each pair, instead of the position of each pair, that can be reversed in different photographs.

To reduce type B variability, a uniform reference will be used, by taking in each photograph a system of orthogonal axes 0XY and applying an orthogonal rotation in such a way that 0X will be parallel to the axis joining pair (1,1′) (Fig. 2). Chromosome No. 1 and its homologue 1′ were chosen because they are easily identifiable and in general they are widely separated in all metaphase plates.

The choice of pair No. 1 as a reference introduced a correlated error in the calculations. However, the study was repeated using as a reference the axis joining pair (5,5′); the results obtained were practically identical to those using the 1,1′ axis.

To fix the orientation of a segment after its orthogonal rotation we chose as

https://doi.org/10.1017/S0016672300019856 Published online by Cambridge University Press



$$\chi = X \cos \alpha + Y \sin \alpha$$

$$\gamma = -X \sin \alpha + Y \cos \alpha$$

Fig. 2. To reduce the variability among the different photographs two orthogonal axes are taken as a reference. The first one corresponds to the segment joining pair No. 1 (1,1′) because it is easily identifiable and its members are usually wide apart in all metaphases.

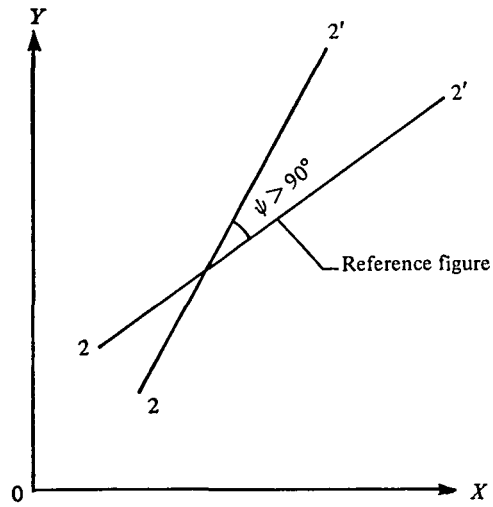


Fig. 3. If the segment joining pair (i,i′) is at an angle $\psi < 90°$ with the same segment in the photograph taken as a reference, its position remains unchanged.

a reference a photograph that contained a complete set of 46 chromosomes. If (i,i') corresponded to a chromosome pair joined by a segment at an angle ϕ with the segment joining the same pair in the reference photograph (Fig. 3), its

https://doi.org/10.1017/S0016672300019856 Published online by Cambridge University Press

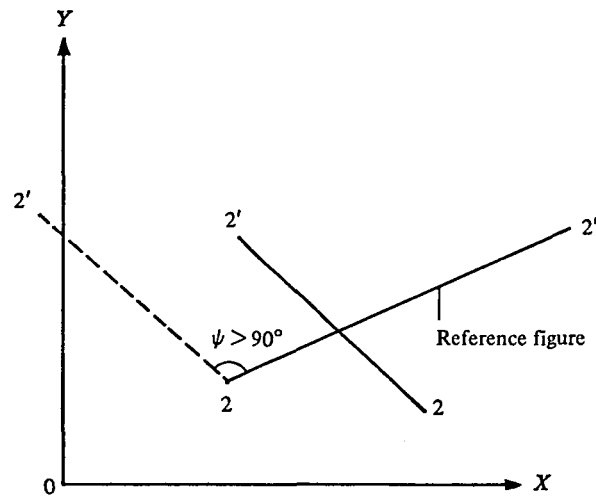


Fig. 4. If the segment joining pair (i,i') is a angle $\psi > 90°$ with the same segment in the reference photograph, its position is reversed.

position remained unchanged when $\psi < 90°$ but was reversed, i.e. $(i,i') \rightarrow (i',i)$ if $90° < \psi < 180°$ (Fig. 4).

Let (x_{ij},y_{ij}) be the coordinates of pair (i,i') on photograph j. The change of coordinates is made in two steps:

$$(x_{ij},y_{ij}) \rightarrow (x_{ij}\cos\alpha + y_{ij}\sin\alpha, \quad -x_{ij}\sin\alpha + y_{ij}\cos\alpha)$$
$$(x'_{ij}) \rightarrow (x'_{ij}\cos\alpha + y'_{ij}\sin\alpha, \quad -x'_{ij}\sin\alpha + y'_{ij}\cos\alpha)$$
$$(x_{ij},y_{ij}), \quad (x'_{ij},y'_{ij}) \rightarrow (x_{ij},y_{ij}), \quad (x'_{ij},y'_{ij}) \quad \text{if} \quad \psi < 90°$$
$$(x_{ij},y_{ij}), \quad (x'_{ij},y'_{ij}) \rightarrow (x'_{ij},y'_{ij}), \quad (x_{ij},y_{ij}) \quad \text{if} \quad \psi > 90°$$

The coordinates (x_{ij},y_{ij}) can be considered as a random sample proceeding from two random variables (X, Y) with a bivariate normal distribution.

(iii) *Statistical models* (See Appendix)

The first model takes into account the causes of variability A and C. It is assumed that the position vector P_{ij} of the mean point of pair (i,i') in photograph j is equal to an ideal fixed vector α_i plus a random deviation. A similar model is also proposed for the position vector joining pair (i,i'). The least squares estimates of the fixed position vectors are then determined.

The second model takes into account the causes of variability A, B and C. It is assumed that the position vector P_{ij} is the sum of a centre of gravity μ (bivariate general mean) plus a systematic deviation α_i due to chromosome i, plus a systematic deviation β_j due to photograph j, plus a random deviation. A formally identical model is proposed for the position vector joining pair (i,i'). The least squares estimates of the parametric vectors μ, α_i and β_j are also determined.

https://doi.org/10.1017/S0016672300019856 Published online by Cambridge University Press

(iv) *The 'centre of symmetry' hypothesis*

The representation of the chromosomes in the metaphase plate would be easier if it were possible to accept the existence of a centre of symmetry for all the chromosomes, except in cases of deviation due to random errors. This hypothesis is formulated by establishing the null hypothesis

$$H_0: \alpha_1 = \alpha_2 = \ldots = \alpha_{23},$$

that is, the position vector of the mean point of pair (i,i') is the same for the 23 chromosome pairs. The verification of this hypothesis must be confirmed by bivariate analysis of the variance in both models.

(v) *Comparison of the segments joining each chromosome pair*

Let $\tilde{\alpha}_i$ be the ideal position vector joining pair (i,i') according to the first model. To be able to represent the chromosomes it is necessary to verify that $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_{23}$ are significantly different, that is, the null hypothesis H_0: $\tilde{\alpha}_1 = \ldots = \tilde{\alpha}_{23}$ can be rejected. This is carried out by bivariate analysis of the variance. If the second model the hypothesis is the same but $\tilde{\alpha}_i$ corresponds to the systematic deviation of chromosome i with respect to the centre of gravity $\tilde{\mu}$.

If furthermore, there is a centre of symmetry, the mean point of each segment will correspond to the origin of coordinates.

(vi) *Canonical representation*

It is preferable to represent the chromosomes in relation to the statistical distance of Mahalanobis (Goodman, 1972) rather than using the Euclidean distance, because the geometrical representation with this distance can be deformed by a correlation effect between variables X and Y. Rao (1952) proposed to represent population means (which in our case would be the mean segments) with respect to canonical axes. The advantages of the procedure are: (1) canonical axes can be represented orthogonally, (2) discrimination among the means is optimal with respect to the distance of Mahalanobis, (3) the error of the mean, which is an elliptic region (in two dimensions) with respect to the original axes, becomes a circular region in the canonical representation.

Rao's method can be used to represent the mean segments $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_{23}$ in the first model. The results of the canonical analysis are the coordinate points (c_{11},c_{12}), $\ldots$, (c_{q1}, c_{q2}) $(q = 23)$ that correspond to one end of the mean segments relative to the other (i.e. one end is taken to be the origin).

The extreme (c_{i1},c_{i2}) is affected by sampling error. It is possible to demonstrate that this extreme is contained within a circular confidence region whose radius depends on the confidence coefficient chosen and from the number of times pair (i,i') is present in the total number of photographs.

Finally, the segment can be translated so that the mid-point is fixed at the centre of symmetry and origin of the coordinates, and the two extremes then have a confidence region around them representing the sampling error (Fig. 5).

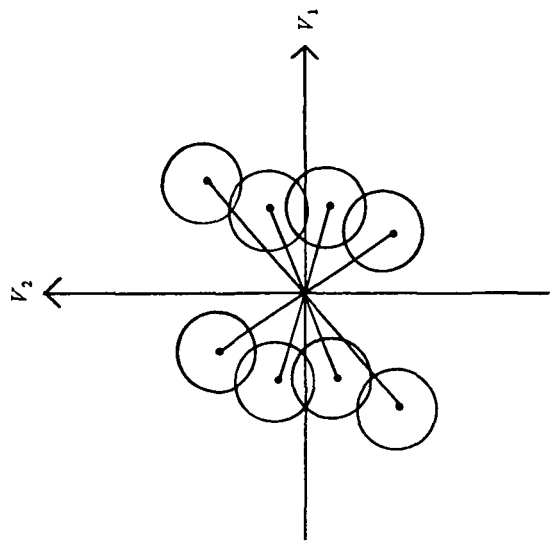


Fig. 5. If there is a centre of symmetry, the segments can be translated, with their mean points coinciding with the origin of the coordinates. The result is the representation of all chromosome pairs with the margin of statistical error corresponding to a confidence circle that depends on the confidence coefficient.

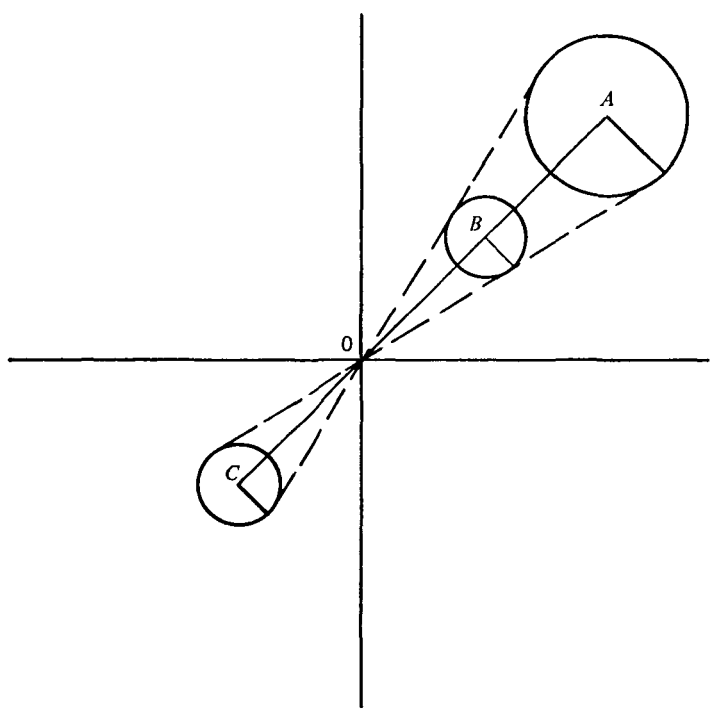


Fig. 6. When a segment with its extremes at 0 and A is translated so that its mean point coincides with 0 and its extremes are at B and C the radius of the confidence region is reduced to half.

https://doi.org/10.1017/S0016672300019856 Published online by Cambridge University Press

The radius of these regions is about half that calculated for the canonical analysis because, when the segments with their origin at the centre of symmetry and their extreme at circle A are translated so that their mean point coincides with the origin of coordinates, the radius of the circle (confidence region) is reduced to half (Fig. 6).

All calculations were carried out twice, using as a reference two different metaphases that contained a complete set of 46 chromosomes.

The second model permits a better control of the variability of the chromosomes, and was applied to the seven metaphases with 46 chromosomes. However, its application to all metaphases studied was impossible for technical reasons, although the program to be used had been established (CANG Program, Cuadras, 1977). Matrix A, used in the least squares estimation of parameters, should have 53 columns and $23 \times 30 - 160 = 530$ lines. Using the CANG program, with the program divided into two parts, the occupation of memory would be 1500 K, which is excessive for the IBM/360 (96 K) of the Laboratorio de Cálculo at the University of Barcelona. Furthermore, to be applied, the CANG program had to be modified by double-precision operations that increased the memory needs.

(3) RESULTS

The comparison of the mean points of the segments joining each chromosome pair using bivariate analysis of variance (Rao, 1973) gave a Fisher-Snedecor F value of 0·383 with 44 and 1028 degrees of freedom which is non-significant and strongly indicates the existence of a centre of symmetry.

The comparison of the segments that join the members of each pair, once fixed at the origin, gave a Fisher-Snedecor F value of 13·3 with 44 and 1028

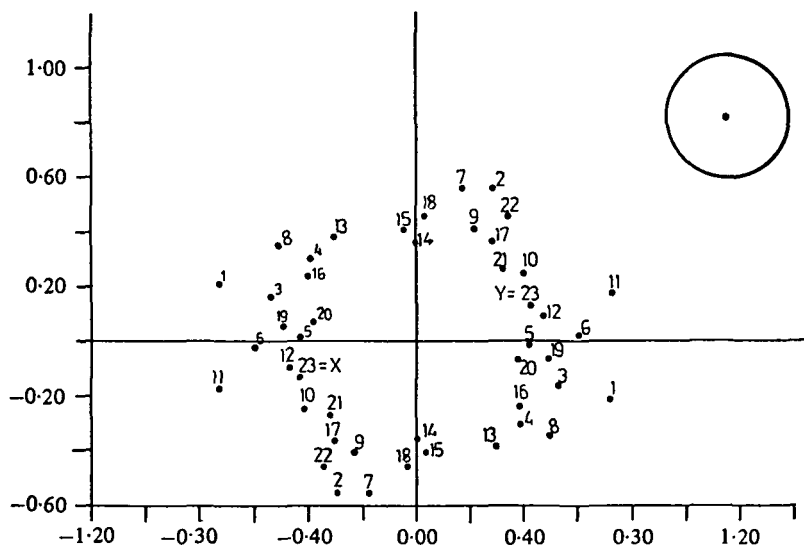


Fig. 7. Distribution of the chromosomes on the metaphase plate and confidence region. Circle of confidence = 90·00%.

https://doi.org/10.1017/S0016672300019856 Published online by Cambridge University Press

degrees of freedom (i.e. $p < 10^{-10}$) which is highly significant and indicates that the chromosomes are not randomly distributed.

The mean position of each segment and the margin of statistical error were determined by generalized canonical analysis (Cuadras, 1974). The mean segments have been represented in relation to the statistical distance of Mahalanobis by taking the reference axes oriented according to a covariance of $-0{\cdot}33952$. Fig. 7 shows the segments joining each chromosome pair with their mean points corresponding to the centre of symmetry and the statistical error of their extremes with a mean radius of 0·4. However, the radius of the circle has been reduced to half because the sampling error of the two extremes of a segment with a fixed mean point is about half the sampling error of one of the extremes when the other is fixed (Fig. 6).

The position of each chromosome pair on the metaphase is also shown in Fig. 7.

(4) DISCUSSION

The results obtained from the application of the statistical model described to a sample of 30 metaphases from 30 normal males indicate that the chromosomes in the metaphase plate are distributed around a centre of symmetry and that the mean point of the segments joining the members of each chromosome pair corresponds approximately to this centre of symmetry. Thus each chromosome mirrors the position of its homologue. The position of the chromosomes is constant within the limits of the confidence region calculated, and the distance of each chromosome to the centre of symmetry is also approximately constant. The confidence regions that do not interfere with each other are statistically different.

The chromosomes show an ordered arrangement on the metaphase plate that probably reflects their distribution in the nucleus. The symmetrical disposition of the homologues is compatible with their association in interphase (Feldman *et al.* 1966; Egozcue, 1968; Fedak & Helgason, 1970 and Yoshida, 1974), as the metaphase corresponds to a flattened nucleus. Their peculiar distribution in metaphase also confirms the absence of proximities between homologues described by Sele *et al.* (1977).

The position of the chromosomes on the metaphase plate does not depend on their size, as already indicated by Sele *et al.* (1977). Some of the larger chromosomes (No. 1, 2) along with some of the smaller ones (No. 18, 22) show a peripheral distribution.

The proximity between pairs No. 14, 15, 21 and 22 and pair No. 9, as well as the relationship between pair No. 13 and pairs No. 1 and 16 could explain the frequent association between the acrocentrics and the heterochromatic regions of these chromosomes (Schmid, Vogel & Krone, 1975). The association of acrocentric chromosomes (satellite association) is probably related, other than to their nucleolar-organizing function, to their spatial proximity, with the possible exception of pair No. 13 which is found a little further away (Schmid & Krone, 1977).

Chromosome No. 14 is the member of the D group closest to pair No. 21, followed by chromosome No. 15. Pair No. 13 shows a different location. This might explain the relatively high frequency of 14/21 translocations in humans, as compared to 15/21 and especially 13/21 translocations (Mikkelsen, 1971).

The proximity between the members of the D and G groups and pair No. 9 and the *Y* chromosome is also in agreement with the observations on the relationship between the heterochromatic region of chromosome 9 and the brightly fluorescent region of the *Y* and the nucleolus (Gagne, Laberge & Tanguay, 1972 and Gagne & Laberge, 1973).

The distribution of chromosomes 9 and 22 is in good agreement with the translocation producing the Philadelphia chromosome (Rowley, 1973) in chronic myelogenous leukaemia.

Finally, the recently described 7/14 translocation in *in vitro* lymphocyte cultures (Welch & Lee, 1978; Beatty-De Sana, Hoggard & Cooledge, 1978; and Hecht *et al.* 1978) shows a good correlation with the position of the chromosome involved.

This statistical technique has been used to study the distribution of the chromosomes in the metaphase plate. It can also be used in systematics (Seal, 1964; Lefebvre, 1976 and Petitpierre & Cuadras, 1977), psychology (Cuadras, Petitpierre & Coll, 1977), medicine (Peris, Romeu & Cuadras, 1975) and paleontology (Martinell & Cuadras, 1977).

(5) APPENDIX

(i) *MODEL I*

The position vector P_{ij} of the mean point of pair (i,i') in photograph j is

$$P_{ij} = \alpha_i + e_{ij},$$

where α_i is an ideal position vector of the mean point of (i,i'), and e_{ij} the random deviation with a bivariate normal distribution.

Parameter estimation

The least squares estimate of $\{\alpha_i\}$ is

$$\underset{1}{\hat{\alpha}} = \left(\frac{1}{n_i}\,\Sigma\tilde{x}_{ij'}\frac{1}{n_i}\,\Sigma\tilde{y}_{ij}\right) = (\hat{\alpha}_{xi'}\hat{\alpha}_{yi}),$$

where n_i = number of times, (i,i') is present on the photograph, and

$$\bar{x}_{ij} = (x_{ij}+x'_{ij})/2; \quad \bar{y}_{ij} = (y_{ij}+y'_{ij})/2.$$

For the position vector that joins pair (i,i') the model is

$$\tilde{P}_{ij} = \tilde{\alpha}_i + \tilde{e}_{ij},$$

where $\tilde{\alpha}_i$ is an ideal position vector, and e_{ij} the random deviation with bivariate normal distribution.

https://doi.org/10.1017/S0016672300019856 Published online by Cambridge University Press

The least squares estimate of $\tilde{\alpha}_i$ is

$$\hat{\tilde{\alpha}}_i = \left(\frac{1}{n_i}\Sigma\tilde{x}_{ij'}\frac{1}{n_i}\Sigma\tilde{y}_{ij}\right),$$

where

$$\tilde{x}_{ij} = x'_{ij} - x_{ij}, \quad \tilde{y}_{ij} = y'_{ij} - y_{ij}$$

The estimates of $\{\hat{\alpha}_i, \hat{\tilde{\alpha}}_i\}$ will correspond to the mean positions of the chromosomes, as will be seen later.

According to the hypothesis,

$$\mathrm{H}_0\colon \alpha_1 = \alpha_2 = \ldots = \alpha_q.$$

This can be tested by bivariate analysis of the variance (Rao, 1973). Let α be the mean position of the centre of symmetry. The estimation of α is

$$\hat{\alpha} = (\hat{\alpha}_x, \hat{\alpha}_y) = \left(\frac{1}{n}\sum_j\sum_i \bar{x}_{ij}, \frac{1}{n}\sum_j\sum_i \bar{y}_{ij}\right),$$

where n = total number of chromosome pairs in all photographs. The following matrixes are found

$$R_1 = \begin{pmatrix} R_1(1,1) & R_1(1,2) \\ R_1(2,1) & R_1(2,2) \end{pmatrix} \qquad R_0 = \begin{pmatrix} R_0(1,1) & R_0(1,2) \\ R_0(2,1) & R_0(2,2) \end{pmatrix},$$

where

$$R_1(1,1) = \sum_j\sum_i (\bar{x}_{ij} - \hat{\alpha}_x)^2 \quad R_1(1,2) = \sum_j\sum_i (\bar{x}_{ij} - \hat{\alpha}_x)(\bar{y}_{ij} - \hat{\alpha}_y) = R_1(2,1),$$

$$R_1(2,2) = \sum_j\sum_i (\bar{y}_{ij} - \hat{\alpha}_y)^2 \quad R_0(1,1) = \sum_j\sum_i (\bar{x}_{ij} - \hat{\alpha}_{xi})^2,$$

$$R_0(1,2) = \sum_j\sum_i (\bar{x}_{ij} - \hat{\alpha}_{xi})(\bar{y}_{ij} - \hat{\alpha}_{yi}) = R_0(2,1) \quad R_0(2,2) = \sum_j\sum_i (\bar{y}_{ij} - \hat{\alpha}_{yi}).$$

Wilk's criterion Λ is

$$\Lambda = \det(R_0)/\det(R_1) \quad 0 < \Lambda \leqslant 1$$

and follows the distribution $\Lambda(2, n-q, q-1)$. Since the number of variables that can be observed is $p = 2$, Λ is equivalent to the F test,

$F = 1 - \sqrt{\Lambda}/\sqrt{\Lambda}/.n-q-1/q-1$ with $2(q-1)$ and $2(n-q-1)$ degrees of freedom.

If F is not significant, one can accept the existence of a centre of symmetry.

Comparison of the segments joining each chromosome pair

Now we establish the linear hypothesis

$$H_0\colon \tilde{\alpha}_1 = \tilde{\alpha}_2 = \ldots = \tilde{\alpha}_q.$$

The procedure is quite similar to the one described for the centre of symmetry hypothesis, but using $\tilde{\alpha}_i$ and $\tilde{x}_{ij}, \tilde{y}_{ij}$. The representation of the chromosomes will be possible if the F test is significant, i.e. if the segments are significantly different. If, furthermore, there is a centre of symmetry, the mean point of each segment will correspond to this centre.

https://doi.org/10.1017/S0016672300019856 Published online by Cambridge University Press

Canonical representation

The method is based on the following steps:

(*a*) The matrix of variances-covariances is determined

$$\Sigma = \begin{pmatrix} \text{var (X)} & \text{cov (X, Y)} \\ \text{cov (X, Y)} & \text{var (X)} \end{pmatrix}$$

and it is estimated using $\hat{\Sigma} = R_0/(n-q)$ (Rao, 1973).

(*b*) Let

$$\psi_i = \hat{\bar{\alpha}}_i = (\hat{\bar{\alpha}}_{xi}, \quad \hat{\bar{\alpha}}_{yi}) = (b_{i1}, b_{i2})$$

$$\bar{b}_1 = \frac{1}{q}\Sigma b_{i1}, \quad \bar{b}_2 = \frac{1}{q}\Sigma\, b_{i2} \quad (q = 23)$$

$$B = \begin{pmatrix} b_{11}-\bar{b}_1; & b_{12}-\bar{b}_2 \\ \cdots\cdots\cdots\cdots & \\ b_{q1}-\bar{b}_1; & b_{q2}-\bar{b}_2 \end{pmatrix}.$$

(*c*) Calculate the matrix $\Sigma_\psi = B^t.B$ and determine the normalized eigenvectors V_1, V_2 of Σ_ψ relative to Σ, i.e.

$$\det(\Sigma_\psi - \lambda\Sigma) = 0, \quad \Sigma_\psi V_i = \lambda_i \Sigma\, V_i \quad i = 1, 2,$$

where λ_1, λ_2 are the eigenvalues of Σ_ψ relative to Σ. The canonical coordinates of segments $\psi_1, \psi_2, \ldots, \psi_q$ referred to the canonical axes V_1, V_2 which are orthogonal (not correlated with respect to Σ) are

$$(c_{i1}, c_{i2}) = (b_{i1}, b_{i2})\, V \quad i = 1, 2, \ldots, q \quad \text{where} \quad V = (V_1, V_2)$$

(*d*) The mean segments are represented by the origin and by points (c_{11}, c_{12}), (c_{21}, c_{22}), ..., (c_{q1}, c_{q2}), that are affected by sampling error. These points are contained within a confidence region, equal to a circle of radius R, depending on the confidence coefficient $1-\epsilon$. For the segment that joins pair (i, i')

$$R_i = \frac{R}{\sqrt{n_i}} \quad \text{with} \quad R_\epsilon^2 = F\epsilon\, 2(n-q)/(n-q-1)$$

with $P(F > F_\epsilon) = \epsilon$, F with 2 and $(n-q-1)$ degrees of freedom.

(ii) *Model II*

This model takes into account the three causes of variability, A, B and C. For pair (i, i') in photograph j the position of the mean joint is

$$P_{ij} = \mu + \alpha_i + \beta_j + e_{ij},$$

where $\mu = (\mu_x, \mu_y)$ = centre of gravity of the positions of all chromosomes (general mean), and $\alpha_i = (\alpha_{xi}, \alpha_{yi})$ = deviation of chromosome i from μ in the ideal model of distribution of all chromosomes, $\beta_j = (\beta_{xj}, \beta_{yj})$ = systematic variation of photograph j, and e_{ij} = random deviation with bivariate normal distribution. We impose the restriction $\sum_i \alpha_i = 0, \quad \sum_j \beta_j = 0.$

https://doi.org/10.1017/S0016672300019856 Published online by Cambridge University Press

The model for the position vector P'_{ij} joining chromosomes (i, i') is formally the same.

The least squares estimates of μ, α_i, β_j and $(\tilde{\mu}, \tilde{\alpha}_i, \tilde{\beta}_j)$ could be calculated without any difficulty if the $q = 23$ chromosome pairs were present in all photographs. However, we assumed that some chromosome pairs or individual chromosomes would be missing or not taken into account in a number of pictures. To solve this problem, it is necessary to build a matrix of design A, to relate the observations to the model (Scheffe, 1959).

The elements a_{ij} of A are 0, 1 or -1 according to the two-way layout without interaction. However, the rows corresponding to the chromosome pairs (i, i') not taken into consideration in a photograph must be omitted. The number of row in A is $n = 23m - f$, where f = total number of pairs not taken into consideration. The number of columns is $23 + m - 1$.

The least squares estimate of the parameters solves the normal equations (t indicates a transposed matrix)

$$(\mu_x, \alpha_{x1}, \ldots, \alpha_{x22}, \beta_{x1}, \ldots, \beta_{xm-1})^t = (A^t.A)^-.A^tY_x,$$

where

$$Y_x = (\bar{x}_{11}, \ldots, \bar{x}_{q1}, \ldots, \bar{x}_{1m}, \ldots, \bar{x}_{qm})^t \quad \bar{x}_{ij} = \frac{x_{ij} + x'_{ij}}{2}.$$

The matrix $(A^t.A)^-$ is a generalized inverse of $A^t.A$ (Pringle & Rayner, 1971) and it can be computed by the algorithm proposed by Golub & Reinsch (1970). The standard inverse of $A^t.A$ does not exist.

The least squares estimate of μ_y, $\alpha_{y1}, \ldots$ is analogous, but replacing $\bar{x}_{ij}$ by $\bar{y}_{ij} = (y_{ij} + y'_{ij})/2$. The estimation of $\tilde{\mu}$, $\tilde{\alpha}_i$, $\tilde{\beta}_j$ is also analogous, but replacing $\bar{x}_{ij}$, $\bar{y}_{ij}$ by $x_{ij} = x'_{ij} - x_{ij}$, $y_{ij} = y'_{ij} - y_{ij}$.

The centre of symmetry hypothesis and the comparison of the segments joining each chromosome pair are tested as follows. In model II the dispersion matrixes R_0 and R_1 must be found according to matrix A (Rao, 1973). Wilks's criterion Λ follows the distribution $\Lambda(2, n-r, q-1)$, where r = rank (A), $q = 23$. Λ is equivalent to the F test

$$F = \frac{1 - \sqrt{\Lambda}}{\sqrt{\Lambda}} \quad \frac{n - r - 1}{q - 1}$$

with $2(q-1)$ and $2(n-r-1)$ degrees of freedom.

The rank of A can be calculated by a numerical procedure. If we compute the singular value decomposition of A, r = rank (A) is then the number of singular values of A different from zero (Golub & Reinsch, 1970).

As in model I, the representation of the chromosomes will be possible if the segments are significantly different. This representation can be carried out by determining the common mean point $\mu + \alpha$ ($\alpha = \alpha_1 = \ldots = \alpha_{23}$ if the centre of symmetry hypothesis is valid) and the segments $\tilde{\mu} + \tilde{\alpha}_1, \ldots, \tilde{\mu} + \tilde{\alpha}_{23}$. However, since we are using statistical estimations of $\mu + \alpha$ and $\tilde{\mu} + \tilde{\alpha}_i$ it is convenient to represent the chromosomes in relation to the distance of Mahalanobis.

Rao's method to represent populations has been extended by Cuadras (1974) to represent functions of parameters in more general models. This technique can be used to represent $\tilde{\mu}+\tilde{\alpha}_1, \ldots, \tilde{\mu}+\tilde{\alpha}_{23}$.

The steps are similar to those in model I, but

(*a*) $\hat{\Sigma} = R_0/(n-r) \quad r = \text{rank}(A)$ (Rao, 1973, p. 546).

(*b*) $\psi_i = \hat{\tilde{\mu}}+\hat{\tilde{\alpha}}_i = (\hat{\tilde{\mu}}_x+\hat{\tilde{\alpha}}_x, \hat{\tilde{\mu}}_y+\hat{\tilde{\alpha}}_y) = (b_{i1}, b_{i2})$.

(*c*) $R_i = R_e(\sum_{h=1}^{n} d^2_{ih})^{\frac{1}{2}} \quad R_e^2 = F\,\dfrac{2(n-r)}{(n-r-1)}$.

$P(F > F_e) = \Sigma$, F with 2 and $(n-r-1)$ degrees of freedom.

$(d_{i1}, \ldots, d_{in})$ are the coefficients of the linear combination of the sampling values that give the least squares estimation of $\tilde{\mu}+\tilde{\alpha}_i$.

REFERENCES

Barton, D. E. & David, F. N. (1962). The analysis of chromosome patterns in normal cell. *Annals of Human Genetics* **25**, 323–329.

Barton, D. E., David, F. N. & Merrington, M. (1965). The relative positions of the chromosomes in the human cell in mitosis. *Annals of Human Genetics* **29**, 139–146.

Beatty-De Sana, J. W., Hoggard, M. J. & Cooledge, J. W. (1978). Ibid. *Nature* **255**, 242–243.

Cohen, M. M., Enis, P. & Pfeifer, C. G. (1972). An investigation of somatic pairing in the muntjak. *Cytogenetics* **11**, 145–152.

Comings, D. E. (1968). The rationale for an ordered arrangement of chromatin in the interphase nucleus. *American Journal of Human Genetics* **20**, 440–460.

Cuadras, C. M. (1974). *Análisis estadístico multivariante y representación canónica de funciones estimables*. Barcelona, Secretariado de Publicaciones de la Universidad de Barcelona.

Cuadras, C. M. (1974). Análisis discriminante de funciones paramétricas estimables. *Trabajos de Estadística e Investigación Operativos* **25**, 3–31.

Cuadras, C. M. (1977). *CANG, versión II: Instrucciones de utilización*. Laboratorio de cálculo, Universidad de Barcelona.

Cuadras, C. M., Petitpierre, E. & Coll, M. D. (1977). Generalized discriminant canonical analysis: applications to biology, psychology, and medicine. *First World Conference on Mathematics at the Service of Man*. Barcelona, Universidad Politécnia de Barcelona.

Egozcue, J. (1968). Spontaneous chromosomal exchanges in Primates. *Genetics* **60**, 771–779.

Fedak, G. & Helgason, S. B. (1970). Somatic association of chromosomes in barley. *Canadian Journal of Genetic Cytology* **12**, 449–500.

Feldman, M., Mello-Sampayo, T. & Sears, E. R. (1966). Somatic association in *Triticum aestivum*. *Proceedings of the National Academy of Sciences* (*Washington*) **56**, 1192–1199.

Gagne, R. & Laberge, C. (1973). Specific cytological segment of number 9 chromosome in man. *Experimental Cell Research* **73**, 239–242.

Gagne, R., Laberge, C. & Tanguay, R. (1972). Interphase association of Y-body' with nucleolus. *Johns Hopkins Medical Journal* **130**, 254–258.

Galperin, H. (1968). Etude de la distribution génerale des 46 chromosomes dans les cellules humaines en metaphase. *Humangenetik* **6**, 118–130.

Galperin, H. (1969). Relative positions of homologous chromosomes or groups in male and female metaphase figures. *Humangenetik* **7**, 265–274.

Golub, G. H. & Reinsch, C. (1970). Singular value decomposition and least square solutions. *Numerische Matematik* **14**, 403–420.

Goodman, M. M. (1972). Distance Analysis in Biology. *Systematic Zoology* **21**, 174–186.

Hecht, F., Mc Caw, B. K., Peakman, D. & Robinson, A. (1978). Ibid. *Nature* **255**, 243–244.

Heneen, W. K. & Nichols, W. W. (1972). Non-random arrangement of metaphase chromosomes in cultured cells of the Indian deer, *Muntiacus muntjak*. *Cytogenetics* **11**, 153–164.

Hoo, Y. Y. & Cramer, H. (1971). On the position of chromosomes in prepared mitotic figures of human fibroblasts. *Humangenetik* **13**, 166–170.

Juricek, D. K. (1975). Non-random chromosome distribution in radial metaphases from Chinese hamster. *Chromosoma* (*Berlin*) **50**, 313–326.

Lefebvre, J. (1976). *Introduction aux Analyses Statistiques Multidimensionals*. Paris: Masson.

Martinell, J. & Cuadras, C. M. (1977). Bioestadística y análisis multivariable aplicados a la comparación de una Poblaciónactual y otra fósil, atribuídos a *Sphaeronassa mutabilis* (*Linné*): aportación a la sistemática del género *Sphaeronassa* Locard, 1866. *Studia Geologica* **13**, 89–103.

Mikkelsen, M. (1971). Identification of G-group anomalies in Down's syndrome by quinacrine dihydrochloride fluorescence staining. *Human Genetics* **12**, 67–73.

Miller, O. J., Mukherjee, B. B., Berg, W. R. & Gamble, A. V. N. (1963*a*). Non-random distribution of chromosomes in metaphase figures from cultured human leukocytes. I. The peripheral location of the Y chromosome. *Cytogenetics* **2**, 1–14.

Miller, O. J., Berg, W. R., Mukerjee, B. B., Gamble, A. V. N. & Christakos, A. C. (1963*b*). Non-random distribution of chromosomes in metaphase figures from cultured human leukocytes. II. The peripheral location of chromosomes 13, 17, 18 and 21. *Cytogenetics* **2**, 152–167.

Moorhead, P. S., Nowell, P. C., Mellman, W. J., Battips, D. M. & Hungerford, D. A. (1960). Chromosome preparations of leukocytes cultured from human peripheral blood. *Experimental Cell Research* **20**, 613–616.

Ockey, C. H. (1969). The position of chromosomes at metaphase in human fibroblasts and their DNA syntesis behaviour. *Chromosoma* (*Berlin*) **27**, 308–320.

Peris, A., Romeu, J. & Cuadras, C. M. (1975). Valoración de la actividad del Ludiomil a través de un ensayo multicéntrico. *Archives of Neurobiology* **38**, 471–484.

Petitpierre, E. & Cuadras, C. M. (1977). The canonical analysis applied to the taxonomy and evolution of the genus *Timarcha* Latr Coleoptera, Chrysomelidae. *Mediterranea* **2**, 13–28.

Pringle, R. M. & Rayner, A. A. (1971). *Generalized Inverse Matrixes with Applications to Statistics*. Griffin's Statistical Monographs and Courses. New York: Hafner.

Rao, C. R. (1952). *Advanced Statistical Methods in Biometric Research*. New York: John Wiley.

Rao, C. R. (1973). *Linear Statistical Inference and its Application*. New York: John Wiley.

Rowley, J. D. (1973). A new, consistent chromosomal abnormality in chronic myelogenous leukemia identified by quinacrine fluorescence and Giesma staining. *Nature* **143**, 290–293.

Schmid, M., Vogel, W. & Krone, W. (1975). Attraction between centric heterochromatin of human chromosomes. *Cytogenetics and Cell Genetics* **15**, 66–80.

Schmid, M. & Krone, W. (1977). On the relationship between the frequency of association and the nuclear constriction of individual acrocentric chromosomes. *Human Genetics* **23**, 267–277.

Scheffe, H. (1959). *The Analysis of Variance*. New York: John Wiley.

Seal, H. L. (1964). *Multivariate Statistical Analysis for Biologists*. London: Methuen.

Sele, B., Jalbert, P., Van Cutsem, B., Lucas, M., Mouriquand, C. & Bouchez, R. (1977). Distribution of human chromosomes on the metaphase plate using banding techniques. *Human Genetics* **39**, 39–61.

Welch, J. P. & Lee, C. L. Y. (1978). Non-random occurrence of 7–14 translocations in human lymphocyte cultures. *Nature* **255**, 241–242.

Yoshida, H. (1974). Somatic association of homologous chromosomes in rye. *Japanese Journal of Genetics* **49**, 45–47.

Yunis, J. J., Roldan, L., Yasmineh, W. G. & Lee. J. C. (1971). Staining of repetitive DNA in metaphase chromosomes. *Nature* **231**, 532–533.