On the Distribution of Phenotypes in XXY Males and their Parents

J. H. EDWARDS

From the Department of Human Genetics, University of Birmingham

Fraser (1963) gave the general algebraic solution of the origin of the extra X chromosome in Klinefelter's syndrome on the implicit assumptions that the non-disjunction was meiotic and that there was no crossing over. If the locus by which the origin of the extra X chromosome is defined is not at the centromere, and if there is crossing over, then, since only the distribution of centromeres is determined by the type of meiotic error, it will not be possible to infer the origin of the centromeres from the phenotypes alone. It is convenient to avoid the word chromosome, and restrict discussion to loci, whose phenotypic effects can be observed, and to centromeres, whose origin is defined by the stage of meiosis at which disjunction failed (Fig. 1).

If, following Fraser, we define by a, b, and c, the proportions of cases in which the extra locus may be inferred to come from the father, or from the mother's second X, or by a double representation of the locus on one maternal X; and we define the proportions of non-disjunctions arising in the testis, when the stage is necessarily at meiosis I, as t, at the first and second meiotic divisions at oogenesis, as u and v, and at a mitotic or somatic division as s; and if the recombination fraction between this locus and the centromere is θ, then

$$a = t \qquad \qquad \qquad (1)$$
$$b = u(1-\theta)+v\theta \qquad (2)$$
$$c = u\theta+v(1-\theta)+s \qquad (3)$$

so that, given estimates of a, b, and c, only limited and conditional inferences can be made of the values of u, v, s, and θ.

Some limited inference is possible for, $b \neq c$ and s is small, or if $b > c$, then clearly $\theta \neq (1-\theta)$ and we can infer that recombinants are appreciably rarer than non-recombinants, either because the locus is near the centromere, or because recombination is impaired.

If we make the further assumptions on the general

grounds of cosmic tidiness that one type of meiotic error at oogenesis is relatively rare and that mitotic errors are also rare, then, on the assumption that deficient chiasma formation is not sufficiently
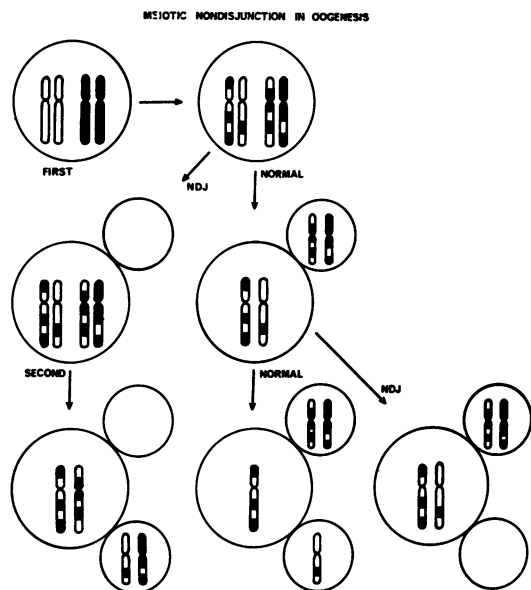

FIG. 1. The consequences of failed disjunction at the first and second meiotic divisions at oogenesis. For diagrammatic convenience the chromosomes in the polar bodies are drawn smaller, and the polar bodies drawn larger. Polar body formation is unlikely if disjunction fails, but the empty bodies are included for simplicity.

associated with non-disjunction to give grossly misleading results, we can regard $b/(b+c)$ or $c/(b+c)$ whichever is smaller as a biassed estimate of the recombination fraction and regard this as a measure of genetic distance commensurate with that derived from pedigree studies.

The most unlikely assumption is that non-disjunction is not predisposed to by a deficiency in chiasma formation: if we regard non-disjunction as a

symptom of chiasma-failure (Bridges, 1916) then the equations reduce to

$$b = u$$
$$\text{and} \quad c = v + s.$$

Fraser calculated the phenotypic expectations of Klinefelter's syndrome for all mating types of dominant diallelic systems, and developed general likelihood functions from the expectations for the whole class of genotypic possibilities (Table I), from which any compound class could be obtained by addition, and from which the proportions of gg Klinefelters may be derived (Table II).

There are advantages in stratifying the analysis

TABLE I

PROPORTIONS OF MATINGS AND WITHIN EACH MATING CLASS, PROPORTIONS OF GENOTYPES EXPECTED WHEN ADDITIONAL LOCUS IS PATERNAL (a), FROM OTHER MATERNAL LOCUS (b), OR THROUGH DUPLICATION OF MATERNAL LOCUS (c)

Father	Mother	Pro-portion	a			b			c		
			GG	Gg	gg	GG	Gg	gg	GG	Gg	gg
G	GG	p^3	1	0	0	1	0	0	1	0	0
G	Gg	$2p^2q$	$\frac{1}{2}$	$\frac{1}{2}$	0	0	1	0	$\frac{1}{2}$	0	$\frac{1}{2}$
G	gg	pq^2	0	1	0	0	0	1	0	0	1
g	GG	p^2q	0	1	0	1	0	0	1	0	0
g	Gg	$2pq^2$	0	$\frac{1}{2}$	$\frac{1}{2}$	0	1	0	$\frac{1}{2}$	0	$\frac{1}{2}$
g	gg	q^3	0	0	1	0	0	1	0	0	1

TABLE II

PROPORTION OF RECESSIVE PHENOTYPES EXPECTED FROM VARIOUS PHENOTYPIC MATING CLASSES

Father	Mother	Proportion g Klinefelters	Proportion when $p = \frac{2}{3}$
G	G	$cq/(1 + q)$	$c/4$
G	g	$b + c$	$b + c$
g	G	$q(a + c)/(1 + q)$	$(a + c)/4$
g	g	$1 \cdot 0$	$1 \cdot 0$
G	?	$cq + bq^2$	$(b + 3c)/9$
G	?	$(1 - pb)/(1 + p)$	$(3 - 2b)/5$
?	G	$q(aq + c)/(1 + q)$	$(a + 3c)/6$
?	g	$1 - ap$	$1 - 2a/3$
?	?	$q(q + cp)$	$(1 + 2c)/9$

TABLE III

RELATIVE PROPORTIONS OF VARIOUS MATING CLASSES AND EXPECTATION OF KLINEFELTERS WHEN $p = \frac{2}{3}$ AS IN Xg^a IN CAUCASIANS

Father	Mother	Pro-portion	a			b			c		
			GG	Gg	gg	GG	Gg	gg	GG	Gg	gg
G	GG	8	8a	0	0	8b	0	0	8c	0	0
G	Gg	8	4a	4a	0	0	8b	0	4c	0	4c
G	gg	2	0	2a	0	0	0	2b	0	0	2c
g	GG	4	0	4a	0	4b	0	0	4c	0	0
g	Gg	4	0	2a	2a	0	4b	0	2c	0	2c
g	gg	1	0	0	a	0	0	b	0	0	c
Total		27									

class by class, and deriving the likelihoods independently from the expectation of each parental mating type. This removes the negative information contributed by sampling vagaries in the proportions of parental mating types sampled and may reduce the false information due to ascertainment. It also makes possible the extraction of information from every mating class in whose offspring manifest allelic variation is possible.

A class-by-class analysis allows some algebraic simplification as only the proportions need to be conserved, but requires electronic computation. However, this is no serious disadvantage as even when information from Klinefelter's syndrome with one or neither parent typed is wasted the algebraic solution is difficult, even with approximations, and the interpretation of the results only rough in the absence of the likelihood distributions due to the poor approximation to normality with the limited numbers available.

This can be greatly simplified in the case of the Xg locus in Caucasians, in whom the dominant allele has a proportion of almost exactly $\frac{2}{3}$ (0·659, Race and Sanger, 1968). If we replace p and q by 2 and 1 in Table I and multiply through, we can obtain the genotypic expectations in integral multiples of a, b, and c, in units of $\frac{1}{27}$ (Table III) from which, by addition, we can define the expectations of phenotypic parental classes, including those with unknown phenotypes (Table II). These expectations, most of which can be simplified further, are simple enough to enumerate on advanced desk machines.
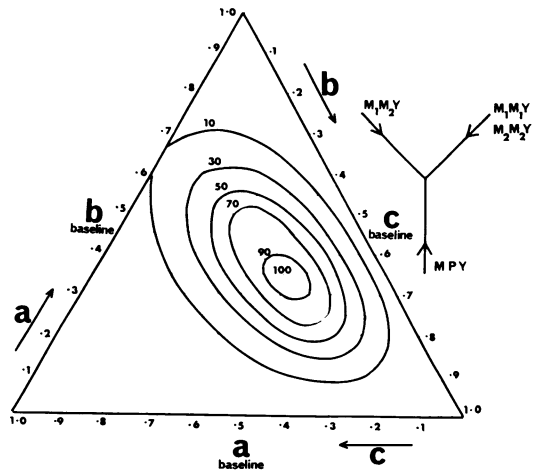

FIG. 2. Likelihood distribution of the three exhaustive and exclusive combinations of paternal and maternal Xg^a loci in XXY Klinefelter's syndrome on the 1966 data. The maximum is standardized at 100 and the ovals are contours of equal relative likelihood.

TABLE IV

LOG LIKELIHOODS (BASE 10, ARBITRARY ABSOLUTE SCALE) FOR TRANSECT THROUGH LOG-LIKELIHOOD SURFACE AT THREE LEVELS OF a (PROPORTION OF PATERNALLY-DERIVED X)

a	Recombination Fraction (θ)									
	0·05	0·10	0·15	0·20	0·25	0·30	0·35	0·40	0·45	0·50
0·30	−11·09	−9·61	−8·87	−8·48	−8·32	−8·33	−8·46	−8·70	−9·03	−9·44
0·40	−11·05	−9·60	−8·85	−8·43	−8·22	−8·14	−8·18	−8·32	−8·53	−8·81
0·50	−11·51	−10·06	−9·29	−8·83	−8·55	−8·40	−8·35	−8·38	−8·47	−8·63

The full information can be summarized by a likelihood surface, in which height is proportional to likelihood, contained within an equilateral triangle, any point of which can represent the proportions a, b, and c as distances from an edge (Fig. 2). This surface was mapped by computing the likelihood of all points for a, b, and c in steps of 0·05, interpolating, and assembling from a series of slabs of balsa wood. The interpretation in terms of likelihoods for various values of θ, of which $c/(b + c)$ is an estimate on the assumption that mitotic non-disjunction does not occur, involves comparing the volumes of a series of thin wedges with their apex at the top of the triangle, and the mid-point of their base at the value of θ specified on the base. This would involve considerable computation. An approximate solution, given below on the more recent data, is to assume the volumes of such wedges are proportional to their ordinate at a horizontal transect near the peak. This is given in Table IV and Fig. 3. This shows a peak at around $\theta = 0·3$ where the likelihood is just over four-fold that
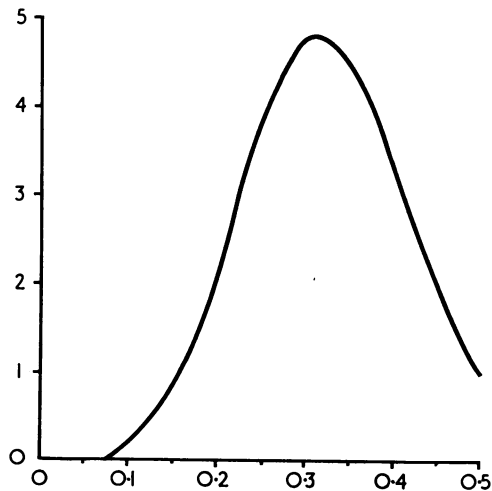
at $\theta = 0·5$. Linkage closer than 0·10 is very unlikely.

An analysis by maximizing likelihoods, class by class, using the classes with both parents of known phenotype, and both of unknown phenotype, was presented in 1966 (Edwards et al, 1966; see also Frøland, 1969). More recent data are given by Race and Sanger (1968), Race (1970), and Sanger, Tippett, and Gavin (1971). The latter data are summarized in Table V, and the expectations derived after maximizing the likelihoods for variations in the parameters a, b, and c, are shown.

FIG. 3. Likelihood transect at $a = 0·4$ which gives approximate likelihood distribution of recombination with the centromere.

TABLE V

DATA FOR XXY KLINEFELTER'S SYNDROME (Sanger et al, 1971) WITH THE EXPECTATIONS DERIVED FROM AN ITERATIVE MAXIMUM LIKELIHOOD SOLUTION. GENE FREQUENCY ASSUMED AT $p = 0·659$

Father	Mother	No.	Observed		Expected		χ^2
			+	−	+	−	
+	+	74	69	5	70·5	3·5	0·73
+	−	10	5	5	3·9	6·1	0·50
−	+	35	31	4	29·9	5·1	0·29
−	−	6	0	6	6·0	0·0	—
+	·	6	3	3	5·3	0·7	9·08
−	·	4	2	2	2·3	1·7	0·71
·	+	42	40	2	38·6	3·4	0·62
·	−	6	2	4	1·5	4·5	0·18
·	·	212	183	29	178·6	33·4	0·68
Total		395	335	60			12·1

Maximum likelihood when $a = 0·39$, $b = 0·43$, $c = 0·18$.

The only serious incongruity between expectation and observation is in the excess of gg Klinefelters whose maternal phenotypes were unknown. When maternal phenotypes were known, the negative excess is very slight (10 observed and 9·6 expected). It is difficult to regard this as other than a fortuitous event. With multiple expectation classes χ^2 is a close approximation even when expectations are small and the effect is to exaggerate the unlikelihood (Cochran, 1954); for 6 degrees of freedom a value of 12·1 is not unduly unlikely ($p = 0·06$). Three or more events would be expected, where the expectation was 0·7, with a probability of 0·03.

One of these cases, whose blood was referred by the author, was exceptional; a son and daughter of his sister died from Duchenne's muscular dystrophy and a plausible case can be made for Lyonization secondary to a deficiency in an X chromosome. If this case is excluded the incongruity is far less ($p = 0.15$).

Since most of the information on Xg^a can be extracted very simply from the parental classes $(+ -)$, $(- +)$, and $(??)$, the last class being the most numerous, simple approximate estimates are possible. Where r_i is the proportion of negative propositi $[g/(G + g)]$ in these three classes, where $i = 1, 2,$ or 3.

$$a' = 1 - r_1$$
$$b' = 1 - 4r_2$$
$$c' = (9r_3 - 1)/2$$
$$s' = a' + b' + c'.$$

From Table II

$$r_1 = 0.50, \quad r_2 = 0.11, \quad r_3 = 0.16$$
$$a' = 0.50, \quad b' = 0.56, \quad c' = 0.22$$
$$a = a'/s' = 0.39$$
$$b = b'/s' = 0.44$$
$$c = c'/s' = 0.17$$

and, assuming all failures of disjunction at the second division, $\theta = c/(b + c) = 0.28$.

The value of c may also be interpreted as the equivalent proportion of males which, when blended with $1 - c$ females, would give the phenotypic proportions found in XXY Klinefelter's syndrome. If the phenotypic proportions of Xg^a positive males and females are p_m and p_f respectively, and the proportion in Klinefelters is p, then

$$p = c \cdot p_m + (1 - c) \cdot p_f$$
$$= c(p_m - p_f) + p_f$$
$$\text{or} \quad c = (p_f - p)/(p_f - p_m)$$
$$\text{or} \quad 1 - c = (p - p_m)/(p_f - p_m).$$

(See also Sanger *et al*, 1971.)

All that can be inferred with certainty at present is: firstly, that if meiotic non-disjunction implies no recombination then both first and second ovarian meioses are common sources of failed disjunctions, the former being commoner; second, if almost all disjunctions are at the first division, then recombination between the Xg locus and the centromere is common, the most likely recombination fraction being less than $\frac{1}{3}$, but the data do not exclude a recombination fraction of 0.5.

Addendum

A likelihood surface does not enclose a volume in a meaningful way, since likelihood is invariant to any transformation of any other axis. The discussion of wedges (p. 436) is a Bayesian argument and should be ignored since there is no substantial prior information.

REFERENCES

Bridges, C. B. (1916). Non-disjunction as a proof of the chromosome theory of heredity. *Genetics*, **1**, 107–163.

Cochran, W. G. (1954). Some methods of strengthening the common χ^2 tests. *Biometrics*, **10**, 417–451.

Edwards, J. H., Ferguson-Smith, M. A., Frøland, A., Linsten, J., Polani, P. E., Race, R. R., and Sanger, R. (1966). The contribution of Xg blood groups of X chromosome aneuploidy. *Third International Congress on Human Genetics*, Abstract 28.

Fraser, G. R. (1963). Parental origin of the sex chromosomes in the XO and XXY karyotypes in man. *Annals of Human Genetics*, **26**, 297–304.

Frøland, A. (1969). *Klinefelter's Syndrome*. Monograph. Costers Bogtrykeri, Copenhagen.

Race, R. R. (1970). The Xg blood groups and sex-chromosome aneuploidy. *Philosophical Transactions of the Royal Society of London*, **259B**, 37–39.

Race, R. R. and Sanger, R. (1968). *Blood Groups in Man*, 5th ed. Blackwell Scientific, Oxford.

Sanger, R., Tippett, P., and Gavin, J. (1971). Xg groups and sex abnormalities in people of Northern European extraction. *Journal of Medical Genetics*, **8**, 417–426.