

Segregation analysis of peripheral neurofibromatosis (NF1)

M Littler, N E Morton

Abstract

Four studies of NF1 support a prevalence of 0.0003 and a carrier incidence at birth of 0.0004. The gene frequency (q) is therefore 0.0002, and the proportion of cases owing to fresh mutation is 0.56. The mutation rate (xq) is 10^{-4} , an unusually high value suggestive of a large gene. Penetrance among subjects examined is virtually complete, and there is no evidence of phenocopies or somatic mutations.

The dominant gene for von Recklinghausen neurofibromatosis (NF1) is located in band 17q11.2. The allele frequency (q) is about 0.0002, and the proportion of cases that are new mutants (x) is close to one half. If there is mutation-selection equilibrium, the mutation rate xq must be about 10^{-4} /locus/generation, an order of magnitude greater than most deleterious mutation rates. Either the locus is unusually large, like Duchenne muscular dystrophy, or the mechanism of mutation is exceptional, or the parameters q and x are substantially in error. This study was undertaken to establish the values of q and x as accurately as possible on present evidence.

Material and methods

We had access to four population based studies: Denmark,¹ Michigan, USA,² Gothenburg, Sweden,³ and Wales.⁴ A person who had two or more of the following was classified as affected: (1) six or more café au lait spots measuring at least 1.5 cm in diameter, or at least 0.5 cm in diameter if the patient was less than 15 years old; (2) axillary freckling; (3) two or more dermal neurofibromas; (4) a plexiform neurofibroma; (5) a first degree relative with NF1 on criteria 1-4. All other persons were classified as

normal if there was an apparently trustworthy informant, and as unknown otherwise. An affected person was taken as a secondary case if ascertained only through a family study, and as a proband otherwise. Probands were ascertained through hospital records, general practitioners, and consultants in each study area.

A handful of probands and their families was excluded because they met the diagnostic criteria of two rare forms: bilateral acoustic neurinoma, which is determined by the NF2 locus on 22q11-q13.1, and segmental neurofibromatosis, which may be a somatic mutation to NF1, sometimes involving the gonads.⁵ Three types of nuclear families were distinguished: (1) index sibships in which one or more children is a proband (multiple incomplete selection); (2) sibships that include an affected parent of a proband or an affected ancestor of a proband (single selection); (3) children of probands and secondary cases (complete selection).

To avoid bias, sibships that could be classified in more than one way were assigned to each class for which they qualified. A total of 2974 subjects from 524 nuclear families was submitted to segregation analysis by the computer programs SEGRAN and POINTER.⁶

Results

If the probability π that an affected person be a proband is constant, the distribution of a probands among r affected sibs in type 1 families will be a truncate binomial (table 1).

Analysed by SEGRAN, the maximum likelihood estimate of π for the pooled data is 0.15 (SD 0.04), but there is significant heterogeneity among studies (table 2). Therefore the prevalence estimate for the pooled data (0.0002) is too small. However, the estimate from the separate studies (0.0004) takes all variation as real and therefore tends to be too high. The mean of 0.0003 is a reasonable compromise and agrees with published reports. This method of estimating prevalence generally works well, even when the ascertainment probability is small,⁷ but it excludes some cases that could be diagnosed by examination but would not be recognised by the typical informant.

CRC Research Group in Genetic Epidemiology, Department of Community Medicine, Southampton General Hospital, Southampton SO9 4XY.

M Littler, N E Morton

Correspondence to Professor Morton.

Table 1 Distribution of probands among affected sibs.

Affected (r)	Probands (a)	Denmark	Michigan	Gothenburg	Wales	Total
2	1	9	17	7	7	40
2	2	1	1	2	3	7
3	1	6	3	3	4	16
3	2	1	0	1	1	3
4	1	1	2	0	0	3
5	1	3	1	0	0	4
6	1	1	0	0	0	1
6	2	0	0	0	1	1
7	1	1	1	0	0	2
Total		23	25	13	16	77

Table 2 Ascertainment and prevalence.

Population	Number (N)	Probands (A)	Ascertainment (π)	Affected (A/ π)	Prevalence (A/ π N)	Residual (χ^2)	χ^2 ($\pi=0.15$)
Denmark	4 000 000	82	0.075	1093	0.00027	3.22	5.03
Michigan	6 070 000	128	0.048	2667	0.00044	1.64	4.30
Gothenburg	338 000	61	0.294	207	0.00061	0.51	1.87
Wales	668 000	73	0.318	230	0.00034	3.56	6.43
Total	11 076 000	344	0.082	4197	0.00038	8.93	17.63
Pooled	11 076 000	344	0.149	2309	0.00021	$\chi^2=8.7$	

Table 3 Segregation analysis.

Class	Sporadic cases		Gene frequency (q)	Penetrance	No of families
	x	SE			
1	0.588	0.039	0.00020	0.766	333
2	0.400	0.115	0.00025	0.599	61
3	0	—	0.00017	0.892	130
1+2	0.560	0.037	0.00020	0.733	394

Table 4 Penetrance by liability.

Class	Liability	Penetrance	No of families	Morbidity risk
1+2	Low	0.574	249	0.00023
1+2	High	0.804	145	0.00032
3	Low	0.644	26	0.00026
3	High	0.951	104	0.00038

Neurofibromatosis poses the difficulty that a relative is more likely to be examined, or medical records retrieved, if the informant suspects affection. Examination status is a poor basis for a liability index, which should be defined independently of phenotype. We therefore first performed segregation analysis with a single liability indicator, assigned a morbidity risk of 0.0003. The data were divided into the three classes defined above. Since parental phenotypes are often in doubt under incomplete selection, parents were taken to be of unknown phenotype for classes 1 and 2. Complete dominance and Mendelian segregation were assumed.

The estimate of sporadic cases (x) is largely determined by index sibships, the pooled estimate by POINTER being 0.56 (SD 0.04) (table 3). Displacement (t) is large, consistent with absence of phenocopies. The gene frequency (q) is 0.0002, corres-

ponding to a carrier incidence at birth of 0.0004, compared with a disease prevalence of 0.0003. The penetrance of 0.73 is the estimated proportion of carriers that would be diagnosed by the mixture of examination and lay informants used in these studies. Many of the 'non-penetrant' cases would be diagnosed if examined now, others would be recognised at a later age, and still others might be diagnosed only by Lisch nodules or other microsymptoms.

To clarify penetrance we classified each family according to whether at least half the members were examined (high liability) or not (low liability), so that all members of a family had the same liability indicator. The same prevalence (0.0003) was assigned to both groups. Under incomplete selection x and q were estimated simultaneously, and under complete selection q was estimated. The four estimates of penetrance ranged from 0.57 to 0.95 (table 4). The

Table 5 Parents of probands and ancestral probands.

Father	Mother	Class 1	Class 2	Total
N	N	178	25	203
N	A	85	20	105
A	N	43	7	50
N	?	3	0	3
?	N	8	0	8
?	?	27	9	36
Total		344	61	405

Table 6 Estimates of the proportion of cases owing to mutation (x).

Source	x	Information
Incomplete selection, parents ignored	0.533	744
Parents of probands	0.567	1458
Pooled	0.556	2202

lowest value is for families under incomplete selection where a minority of members was examined. The unexamined members were often classified by relatives living apart. The highest penetrance is for families under complete selection where a majority of members was examined. The unexamined members were usually classified by a parent. Estimates of penetrance for examined persons approach 100%.² It is likely that lower estimates of penetrance reflect partial examination rather than incomplete expression at first appearance.⁸

Estimates of penetrance were multiplied by incidence at birth (0.0004) to give morbid risk. At these values the gene frequency is 0.0002 and the proportion of sporadic cases is 0.533 (SD 0.037).

Parents of probands and ancestral probands provide additional information about x (table 5). Among the mating types with both parents classified, the proportion Q with an affected parent is $P(1-x)$, where P is penetrance in parents, and a sample of N independent probands gives information $NP^2/Q(1-Q)$ about x . Segregation analysis does not give direct evidence about P . Since a parent with inadequate information tends to be classified as of unknown phenotype, penetrance among classified parents plausibly approaches 1. At that value the estimate of x is 0.567 (SD 0.026). The values from segregation and parents are in good agreement, and the pooled estimate is 0.556 (SD 0.021) (table 6).

Discussion

Segregation analysis of NF1 is complicated by variable expression, indefinite age of onset, and incomplete examination. Nevertheless, estimates of the basic parameters in different populations and by different methods have been remarkably consistent. Borberg¹ observed that 84 probands had 60 children. The average age of probands was 38.5 if living and 36.9 if dead. He attributed the reduced fertility to

celibacy induced by disfigurement. The 60 children were distributed over 22 probands, indicating that reproduction after marriage was in the normal range. Borberg did not estimate the mutation rate, because he considered that variable expressivity of carriers precluded an estimate of gene frequency.

Crowe *et al*² were able to carry the analysis further. They estimated the frequency of neurofibromatosis in the state of Michigan as one case in every 2500 to 3300 persons. Although they did not distinguish this prevalence from the incidence of carriers at birth, we may reasonably equate incidence to their upper estimate of prevalence, from which the gene frequency is exactly the same as our estimate. Without allowing explicitly for incomplete ascertainment they estimated the frequency of sporadic cases to be between 0.426 and 0.642, with a mean of 0.534. Independently they estimated a selection coefficient of 0.473, which at mutation-selection equilibrium should equal the frequency of sporadic cases. Although failure to marry was the major selective factor, fertility among married probands was estimated to be only 0.795 relative to sib controls.

Samuelsson³ found that the prevalence of NF1 in Gothenburg peaked for age 40 to 50 at 0.0003, suggesting reduced life expectancy. Allowing for incomplete ascertainment and penetrance, this is consistent with a carrier frequency at birth of 0.0004, and therefore the same gene frequency we have estimated. She observed 29 sporadic cases among 74 patients, taking a case as sporadic only if there was reliable negative evidence about both parents. Since the same condition was not imposed on familial cases, whose diagnosis is easier, this must tend to give an underestimate of sporadic cases.

Finally, Huson *et al*⁹ estimated the prevalence of NF1 in children to be 0.0004. Among 93 probands there were 41 apparently sporadic cases. Relative fitness was estimated by the method of Tanaka¹⁰ as the ratio of the affection frequency among parents and offspring of index cases. The complement of this (0.53) is the selection coefficient. This simple calculation is complicated by unexamined relatives and age dependent penetrance, but it agrees well with other methods.

It therefore appears that the gene frequency at birth is 0.0002, and half of these genes are lost through selection and replaced by mutation at the rate of 10^{-4} /locus/generation. This is an unusually high frequency, equalling the gene for Duchenne muscular dystrophy. Perhaps the NF1 locus is uncommonly large. Alternatively, the NF1 mutant may arise by transduction, unequal crossing over, or other exceptional mechanisms. Closely linked RFLPs can establish whether mutation occurs preferentially in sperm, as is the case for some other dominant genes. An increased mutation rate with paternal age has been claimed,¹¹ but was not found in the studies analysed

here. Convincing evidence must await probes to the NF1 locus, which alone are capable of identifying the mutated gamete in parent-offspring pairs. Using closely linked probes, Jadayel *et al*¹² have recently reported that 12 of 14 mutants were paternal in origin and not associated with paternal age or crossing over.

Crowe *et al*² estimated fitness from reproductive performance, while Huson *et al*⁹ used the Tanaka method based on affection rates. Both approaches indicated a substantial loss of fitness, markedly greater in males than females. Since the population frequencies are nearly equal at all ages, differential fitness must reflect lower marriage rates for male carriers. The rates of selective mortality and fertility of married carriers remain to be established. Close agreement of direct and indirect estimates of mutation rate suggests that few sporadic cases can be somatic mutants or other mosaics in which the gonad is spared.

1 Borberg A. Clinical and genetic investigations in tuberous sclerosis and Recklinghausen's neurofibromatosis. *Acta Psychiatr Neurol Scand* 1951;suppl 71:1-239.

- 2 Crowe FW, Schull WJ, Neel JV. *A clinical, pathological, and genetic study of multiple neurofibromatosis*. Springfield, Ill: Charles C Thomas, 1956.
- 3 Samuelsson B. *Neurofibromatosis (v Recklinghausen's disease). A clinical, psychiatric and genetic study*. Thesis, University of Göteborg, 1981.
- 4 Huson SM, Harper PS, Compston DAS. von Recklinghausen neurofibromatosis. A clinical and population study in South-East Wales. *Brain* 1988;111:1355-81.
- 5 Jung EG. Segmental neurofibromatosis (NF-5). *Neurofibromatosis* 1988;1:306-11.
- 6 Morton NE, Rao DC, Lalouel JM. *Methods in genetic epidemiology*. New York: Karger, 1983.
- 7 Barrai I, Mi MP, Morton NE, Yasuda N. Estimation of prevalence under incomplete selection. *Am J Hum Genet* 1965;17:221-36.
- 8 Riccardi VM, Lewis RA. Penetrance of von Recklinghausen neurofibromatosis: a distinction between predecessors and descendants. *Am J Hum Genet* 1988;42:284-9.
- 9 Huson SM, Compston DAS, Clark P, Harper PS. A genetic study of von Recklinghausen neurofibromatosis in south east Wales. I Prevalence, fitness, mutation rate, and effect of parental transmission on severity. *J Med Genet* 1989;26:704-11.
- 10 Tanaka K. A new simplified method for estimating relative fitness in man. *Jinrui Idengaku Zasshi* 1974;19:195-202.
- 11 Risch N, Reich EW, Wishnick MM, McCarthy JG. Spontaneous mutation and parental age in humans. *Am J Hum Genet* 1988;41:218-48.
- 12 Jadayel D, Fain P, Upadhyaya M, *et al*. Paternal origin of new mutations in Von Recklinghausen neurofibromatosis. *Nature* 1990;343:558-9.