Modelling the Effects of Penetrance and Family Size on Rates of Sporadic and Familial Disease

Ammar Al-Chalabi^a Cathryn M. Lewis^{b, c}

^aDepartment of Clinical Neuroscience, Medical Research Council Centre for Neurodegeneration Research,

^bMedical Research Council Social, Genetic and Developmental Psychiatry Centre, and ^cDepartment of Medical and Molecular Genetics, King's College London, London, UK

Key Words

Amyotrophic lateral sclerosis • Familial • Genetic models • Penetrance • Population genetics • Segregation analysis • Sporadic

Abstract

Background/Aims: Many complex diseases show a diversity of inheritance patterns ranging from familial disease, manifesting with autosomal dominant inheritance, through to simplex families in which only one person is affected, manifesting as apparently sporadic disease. The role of ascertainment bias in generating apparent patterns of inheritance is often overlooked. We therefore explored the role of two key parameters that influence ascertainment, penetrance and family size, in rates of observed familiarity. **Methods:** We develop a mathematical model of familiarity of disease, with parameters for penetrance, mutation frequency and family size, and test this in a complex disease: amyotrophic lateral sclerosis. **Results:** Monogenic, high-penetrance variants can explain patterns of inheritance in complex diseases and account for a large proportion of those with no apparent family history. With current demographic trends, rates of familiarity will drop further. For example, a variant with penetrance 0.5 will cause apparently sporadic disease in 12% of

families of size 10, but 80% of families of size 1. A variant with penetrance 0.9 has only an 11% chance of appearing sporadic in families of a size similar to those of Ireland in the past, compared with 57% in one-child families like many in China.

Conclusions: These findings have implications for genetic counselling, disease classification and the design of gene-hunting studies. The distinction between familial and apparently sporadic disease should be considered artificial.

Copyright © 2011 S. Karger AG, Basel

Introduction

Complex diseases are those in which the inheritance pattern is not clearly understood. Many complex diseases show a diversity of inheritance patterns ranging from familial disease, manifesting with autosomal dominant inheritance, through to simplex families in which only 1 person is affected, manifesting as apparently sporadic disease [1]. It is usually assumed that large-effect rare variants cause the familial disease [2] and (through the liability threshold model) a polygenic contribution of small-effect variants or low-penetrance variants is responsible for the simplex cases [3]. The role of ascertainment bias in generating apparent patterns of inheritance

is often overlooked, and with changes to population demographics, such bias plays a crucial role in whether a disease appears familial or sporadic. This has implications for genetic counselling, strategies for gene-hunting and disease classification.

For a disease to appear familial, several social and clinical factors must come together. First, the family needs to be of sufficient size that more than one person could be affected. Second, the disease phenotype needs to be consistent and straightforward enough to be diagnosed reliably. Third, the disease needs to be one that is not associated with undue stigma so that it is not hidden from other family members. Fourth, the proband needs to be in contact with their family. Fifth, if the disease affects people later in life, family members need to have reached the age of risk. Unless all these criteria are met, even a genetic disease caused by a high-penetrance mutation will not appear familial. With the trend to reduction in family sizes and increased frequency of family breakdown, two of these factors are negatively affected. If the disease is late onset and rare, has diagnostic mimics like Alzheimer's disease, or has phenocopies like cancers or coronary artery disease, the situation is compounded.

Many diseases can illustrate this problem. For example, localisation of the *BRCA1* gene (MIM 113705) in which mutations cause breast and ovarian cancer, required large families, such as those ascertained from the Utah Population Database [4, 5]. In other families segregating *BRCA1* mutations, linkage was only clear because of large sibships and a high ratio of female-to-male offspring. Without these properties, it would have been more difficult to be certain that the disease was truly familial, particularly as phenocopies are possible [6]. If this concept is true for high-penetrance mutations in genes such as *BRCA1* (~65% penetrance), it will be even more important for lower-penetrance mutations.

Another disease that can illustrate the problem well is amyotrophic lateral sclerosis (ALS; MIM 105400). This is a disease with complex inheritance with a lifetime risk of about 1 in 300 and onset in later life, typically after the age of 50 [7, 8]. Between 5 and 10% of affected individuals have a family history, often consistent with autosomal dominant inheritance, and the familial form of the disease is indistinguishable from the apparently sporadic form both clinically and pathologically [9, 10]. The disease is straightforward to diagnose, despite the variable phenotype and lack of a diagnostic test, although previous generations may have been given the diagnosis of 'wasting disease', multiple sclerosis or muscle disease [11]. ALS is the commonest reason to seek euthanasia and the

median survival is two years, so it is greatly feared and family members may not discuss it.

Mutations in several genes have now been found in familial ALS, but in every case similar mutations have been identified in a proportion of those with apparently sporadic ALS (<http://alsod.iop.kcl.ac.uk> [12]). The commonest cause of familial ALS is mutation in the *SOD1* gene (OMIM 147450), accounting for about 20% of familial cases [13]. *SOD1* mutations are generally high-penetrance, autosomal dominant mutations, with typically 90% penetrance by age 70 [14]. Between 2 and 7% of those with apparently sporadic ALS also harbour *SOD1* mutations [15]. Even if we take the lower estimate of the frequency of *SOD1* mutations in those with apparently sporadic ALS, an individual with *SOD1* mutation is just as likely to have apparently sporadic ALS as familial ALS despite the high penetrance of *SOD1* mutations. This is intuitively surprising, but replicated across many different diseases. These considerations have an impact on the design of next-generation sequencing studies, in selecting cases for exon or whole-genome sequencing.

We show here that the presence of a rare, moderate- or high-penetrance variant is compatible with many observed familial and apparently sporadic patterns of disease, that the rate of apparently sporadic disease will increase as population demographics change, and that the distinction between familial and apparently sporadic disease should be considered artificial. We develop a mathematical model of familiarity of disease, with parameters for penetrance, mutation frequency and family size. This is applied to ALS.

Materials and Methods

Model

We assume a rare dominant mutation is necessary but not sufficient for disease to occur, so that penetrance is incomplete, but no non-mutation carriers are affected. We assume that an affected individual is ascertained with probability 1. Since the mutation is rare, homozygosity is not included in the model and exactly one parent in a family will be a mutation carrier. We also assume no de novo mutations. For nuclear families in which a parent carries such a variant, there are three possible states for the family: unascertained (no one in the family is affected), simplex (1 person is affected) and familial (>1 family members are affected).

We first derive some basic probabilities. The disease genotype is *Dd*, penetrance *f*, and so the probability an individual is affected (*A*) is

$$P(A | Dd) = f, \text{ and } P(A | dd) = 0.$$

In a nuclear family segregating a mutation, let *a* (= 0, 1) be the number of parents affected and *c* (= 0, 1, 2, ...) the number of off-

spring affected. For a sibship of size N , the number of affected siblings has a binomial distribution, $\text{Bin}(N, f/2)$, assuming the mutation is transmitted with Mendelian probability of 0.5 from the mutation carrier parent to each sibling. Thus,

$$P(c=0) = \left(1 - \frac{f}{2}\right)^N,$$

and

$$P(c=1) = N \left(\frac{f}{2}\right) \left(1 - \frac{f}{2}\right)^{N-1},$$

so

$$P(c \geq 2) = 1 - \left[\left(1 - \frac{f}{2}\right)^N + N \left(\frac{f}{2}\right) \left(1 - \frac{f}{2}\right)^{N-1} \right]$$

The non-mutation carrier parent will be unaffected. The probability the mutation carrier parent will be unaffected is $P(a=0) = 1 - f$.

From this we can derive the probability that no family member is affected and the family is therefore unascertained. Conditioning on one parent being a mutation carrier, so that the parental mating type is $Dd \times dd$,

$$\begin{aligned} P(\text{family unascertained} | Dd \times dd) \\ &= P(a=0 | Dd \times dd) P(c=0 | Dd \times dd) \\ &= (1-f) \left(1 - \frac{f}{2}\right)^N \end{aligned}$$

The disease will appear sporadic (family is simplex) if either the mutation carrier parent is affected or exactly 1 child is affected, but not both.

$$\begin{aligned} P(\text{family is simplex} | Dd \times dd) \\ &= P(a=1 | Dd \times dd) P(c=0 | Dd \times dd) + \\ &\quad P(c=1 | Dd \times dd) P(a=0 | Dd \times dd) \\ &= f \left[\left(1 - \frac{f}{2}\right)^N + N \left(\frac{f}{2}\right) \left(1 - \frac{f}{2}\right)^{N-1} (1-f) \right] \end{aligned}$$

Since the probability of a family being unascertained, simplex or familial must sum to 1, the familial rate is:

$$\begin{aligned} P(\text{familial}) &= 1 - P(\text{unascertained}) - P(\text{simplex}) \\ &= 1 - \left[(1-f) \left(1 - \frac{f}{2}\right)^N + f \left(1 - \frac{f}{2}\right)^N + \left[N \left(\frac{f}{2}\right) \left(1 - \frac{f}{2}\right)^{N-1} (1-f) \right] \right] \\ &= 1 - \left[\left(1 - \frac{f}{2}\right)^N + N \left(\frac{f}{2}\right) \left(1 - \frac{f}{2}\right)^{N-1} (1-f) \right] \end{aligned}$$

In reality, we cannot detect the unascertained families. We define the familial rate as the reported proportion of ascertained families with familial disease, which is given by:

$$\frac{P(\text{familial})}{P(\text{familial}) + P(\text{simplex})}$$

and the reported simplex rate as:

$$\frac{P(\text{simplex})}{P(\text{familial}) + P(\text{simplex})}$$

Results

We now explore the relationships between the model parameter values of genetic penetrance (f), sibship size (N), and the familiarity of a trait. This model enables us to determine population level disease probabilities. Firstly, focussing on a mutation, we calculate the probability that a family segregating the mutation remains unascertained, is simplex or has multiple affected individuals and the disease therefore appears familial. Secondly, focussing on cases carrying the mutation, we calculate the probability that the case appears sporadic or familial.

The probability of a nuclear family being unascertained, simplex or familial given a mutation penetrance f is shown in figure 1, for $N = 1, 2, 3$ offspring per family. For penetrance 0.5, the probability that a family has at least 2 cases of disease increases from 0.125 with a single sibling to 0.25 with 2 siblings, and to 0.367 for 3 siblings. There is a concomitant decrease in probability that the family is unascertained (with no family members affected) from 0.375 to 0.281 to 0.211 for 1, 2, and 3 siblings.

These figures can also be used to estimate the probability that an observed case is familial (with a first-degree relative affected) or apparently sporadic. The probability that a family with a rare dominant mutation is simplex decreases approximately linearly with penetrance for small family sizes ($N = 1, 2$), but has a much sharper decrease for large family sizes (fig. 2).

These calculations may be used to guide next-generation sequencing studies. A common design is to perform exome sequencing of familial cases of a genetic disorder, because these are thought likely to carry a rare mutation. For an individual from a nuclear family to be sequenced, the disorder will need to be present in at least 2 members. A mutation-carrying nuclear family has a 60% probability of showing familial disease provided the penetrance is at least 71% with 3 siblings, but requires penetrance of 85% if the family has only 2 siblings. Even with full penetrance, the probability that such a family with 2 siblings is identified as familial is only 75% (because the mutation carrier parent transmits a wild-type chromosome to both siblings with probability one quarter).

Case Study 1: *SOD1* Mutation in ALS

We will take the overall familial rate of ALS as 10%, with 20% of familial cases due to *SOD1* mutation, while 2% of those with apparently sporadic disease will be assumed to have *SOD1* mutation. Thus, *SOD1*-mediated ALS comprises 4% of the total, being familial in 50% and apparently sporadic in 50% of those with *SOD1* mutation.

Fig. 1. Effect of mutation penetrance and family size on ascertainment and classification. Probability of a nuclear family being unascertained (yellow), simplex (orange) or familial (red) given a mutation penetrance f , for $N = 1, 2, 3$ offspring per family.

Several common *SOD1* mutations have high penetrance, typically of the order of 90% by age 70 [14]. Many are lower penetrance.

Assuming an average penetrance for *SOD1* mutations of 0.7 in a purely genetic model, with a typical modern family size of 2 parents and 2 children, our model predicts that about 13% of *SOD1* mutations do not result in ascertained disease. Of the remainder, 50% would be in apparently sporadic ALS and 50% in familial, giving rates similar to those observed.

Case Study 2: Chromosome 9p21.2 Locus in ALS

Some complex diseases have a locus both showing linkage in families and association in those with apparently sporadic disease. The problem then becomes differentiating between two scenarios. Either the genetic lesion is the same in those with familial and apparently sporadic disease, or there is a large-effect rare variant in the families and polygenic or small-effect variant in those with apparently sporadic disease. This problem exists in a form of ALS with frontotemporal dementia, which is linked to chromosome 9p21.2 in 7 different families and shows independently replicated association to the same locus in several populations [16–25].

From a large 9p21.2-linked family with a sibship of 16, the penetrance of the variant is estimated to be approximately 0.2 [17]. Our model shows a rare variant with this penetrance will result in just 65% of similar large families appearing to have familial disease. 15% of such families will have no ALS cases and remain undetected despite the moderately high penetrance and large number of family members. Assuming families with a more typical sibship size of 2, in 87% of cases resulting from this variant, the disease will appear sporadic. This means that for most families the variant is nearly 7 times more likely to result in apparently sporadic than familial disease. If we assume about 13% of familial cases are a result of mutation at this locus, about 9% of apparently sporadic cases would be, without needing to invoke a different, smaller-effect variant at this locus.

Case Study 3: Effect of Differing Family Sizes on Familiarity Rates

Average family sizes have been decreasing across many countries. For women now aged 40–44 years, the average number of offspring in England is 1.9 [26] and in Italy 1.4 (http://www.unicef.org/infobycountry/italy_statistics.html). In older age groups the equivalent sizes are 2.2 for

Fig. 2. Effect of family size on observed familial and apparent sporadic rates under a pure genetic model. Purple = Family size of 10; red = family size of 4; blue = family size of 3; black = family size of 2; green = family size of 1.

Fig. 3. Effect of differing family sizes by generation on ascertainment and classification. Probability of a nuclear family being unascertained (yellow), simplex (orange) or familial (red) by family size for women born in England and Wales in 1937 and 1964. Dashed blue lines on 1964 birth cohort show equivalent lines from 1937 cohort.

England and 2.5 for Italy. This is also true in countries with traditionally large families such as Ireland, where the average number of offspring has decreased from 3.9 to 2.0 (http://www.unicef.org/infobycountry/ireland_statistics.html). Some countries such as China enforce a one-child policy [27]. We therefore model the effect of decreasing family sizes on observed familiarity rates.

As can be seen from figure 2, the effect of family size on observed rates of familiarity is particularly large for intermediate penetrance variants. A variant with penetrance 0.5 will cause apparently sporadic disease just 12% of the time in populations with an average of 10 offspring per family, but 80% of the time in populations with an average of 1 offspring per family.

Even a variant with penetrance as high as 0.9 would be vulnerable to this ascertainment bias, with only an 11% chance of appearing sporadic in families of a size similar to those of Ireland in the past, compared with a 57% chance of appearing sporadic in one-child families like many in China.

Case Study 4: Effect of Population Demographic Changes on Familiarity Rates in a Defined Geographical Region

In England and Wales, the mean number of offspring has decreased from 2.4 for women born in 1937 to 1.9 for women born in 1964, the last year for which completed family statistics are available [26]. The distribution of sib-

ship sizes (0, 1, 2, 3, ≥ 4 sibs) in 1937 and 1964 was used to calculate the probability that a rare dominant mutation is ascertained or results in apparently sporadic or familial cases (fig. 3). The results show that changes in birth rate across this time period have had little effect on the probability that a case is apparently sporadic or familial. For example, for a mutation with 50% penetrance, the probability that a family segregating this mutation is unascertained increased slightly from 25.3 to 26.1% across this time period. The probability that a case is familial showed little difference between the two birth cohorts at any penetrance, with decreases of $<2\%$ from 1937 to 1964.

Discussion

By modelling the effect of penetrance on rates of familiarity, we have shown that even monogenic, high-penetrance variants can explain patterns of inheritance in complex diseases and account for a large proportion of those with no apparent family history. With current demographic trends, rates of familiarity will drop further.

Previous studies have explored the proportion of apparently sporadic cases for polygenic diseases [3] and the effects of ascertainment bias on pedigree sampling and parameter estimates in linkage studies of complex disease [28, 29]. For many complex diseases, familiarity rates are of the order of 5 to 10%. Previous studies have shown that under the assumption of polygenic inheritance, the number of predicted apparently sporadic cases is high [3]. If such diseases were in fact monogenic Mendelian traits caused by disease genes with a penetrance of 0.05 to 0.2, they would also have the observed distribution of familial and simplex families across common nuclear family sizes of from 1 to 4 children. Genome-wide association studies have failed to explain much of the heritability of complex diseases [30]. One explanation is that rare variants of moderate or large effect that might not be easily assayed by tag SNPs using current statistical methods are responsible for at least some of the 'missing heritability', and our findings support this possibility. Given that there are about 20,000 genes but more than 20,000 traits, simultaneous additive effects, pleiotropy, epistasis and other factors such as structural variation and modification of DNA and RNA must also contribute [31].

If a significant proportion of apparently sporadic disease is caused by moderate- to high-penetrance rare variants, this has implications for genetic counselling, traditionally confined to those with a family history of disease. As family sizes reduce, many of those without a family

history will nonetheless carry a monogenic form that will pass on to their offspring. The converse is that for diseases in which family history might be regarded as a risk factor, the reduction in family sizes means that this will become unreliable unless the family size is also taken into account.

Large families are rarer now, with 10% of women born in 1964 having families of size 4 or more, compared with 20% of those born in 1937 [26], and as can be seen from figure 2, it is these larger families that provide much of the information on familial disease for intermediate penetrance mutations. However, because the commonest family size was 2 for both cohorts, the effect on average family size is small and not sufficient to significantly impact familiarity rates. As demographic trends continue, and particularly in countries with an enforced small average family size, rates of familiarity will reduce significantly.

Next-generation sequencing study designs often focus on those with familial disease, but for intermediate and lower penetrance mutations particularly, it would also make sense to include apparently sporadic cases, particularly those from small families in which the apparently sporadic nature may well be an artifact of ascertainment.

We have shown that a proportion of apparently sporadic disease will inevitably be a result of high-penetrance monogenic variants. Conversely, a disease variant for a polygenic trait with, for example, a penetrance of just 0.005 will still occur in a finite number of families, even disregarding the probability that there will be contributions from other genes and the environment to risk. Thus, the distinction between familial and apparently sporadic disease is artificial. Although it guides gene hunting methods and thus could be considered useful, as family sizes fall, unless there is a founder effect for at least one disease variant, it will become more difficult to detect genes using tag SNPs because of the shift of individuals with rare monogenic variants from the familial pool to the apparently sporadic pool.

Acknowledgements

We thank the NIHR specialist Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust (SLaM) and the Institute of Psychiatry, King's College London. The research leading to these results has received funding from the European Community's Health Seventh Framework Programme (FP7/2007–2013) under grant agreement No. 259867, and the Actuarial Profession.

References

- Bodmer W, Bonilla C: Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet* 2008;40:695–701.
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttman AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM: Finding the missing heritability of complex diseases. *Nature* 2009;461:747–753.
- Yang J, Visscher PM, Wray NR: Sporadic cases are the norm for complex disease. *Eur J Hum Genet* 2010;18:1039–1043.
- Goldgar DE, Cannon-Albright LA, Oliphant A, Ward JH, Linker G, Swensen J, Tran TD, Fields P, Uharriet P, Skolnick MH: Chromosome 17q linkage studies of 18 Utah breast cancer kindreds. *Am J Hum Genet* 1993;52:743–748.
- Goldgar DE, Fields P, Lewis CM, Tran TD, Cannon-Albright LA, Ward JH, Swensen J, Skolnick MH: A large kindred with 17q-linked breast and ovarian cancer: genetic, phenotypic, and genealogical analysis. *J Natl Cancer Inst* 1994;86:200–209.
- Bowcock AM, Anderson LA, Friedman LS, Black DM, Osborne-Lawrence S, Rowell SE, Hall JM, Solomon E, King MC: THRA1 and D17S183 flank an interval of <4 cM for the breast-ovarian cancer gene (BRCA1) on chromosome 17q21. *Am J Hum Genet* 1993;52:718–722.
- Beleza-Meireles A, Al-Chalabi A: Genetic studies of amyotrophic lateral sclerosis: controversies and perspectives. *Amyotroph Lateral Scler* 2009;10:1–14.
- Johnston CA, Stanton BR, Turner MR, Gray R, Blunt AH, Butt D, Ampom MA, Shaw CE, Leigh PN, Al-Chalabi A: Amyotrophic lateral sclerosis in an urban setting: a population based study of inner city London. *J Neurol* 2006;253:1642–1643.
- Andersen PM, Nilsson P, Keranen ML, Forsgren L, Hagglund J, Karlsborg M, Ronnevi LO, Gredal O, Marklund SL: Phenotypic heterogeneity in motor neuron disease patients with Cu/Zn-superoxide dismutase mutations in Scandinavia. *Brain* 1997;120:1723–1737.
- Byrne S, Walsh C, Lynch C, Bede P, Elamin M, Kenna K, McLaughlin R, Hardiman O: Rate of familial amyotrophic lateral sclerosis: a systematic review and meta-analysis. *J Neurol Neurosurg Psychiatry* 2011;82:623–627.
- Kurian KM, Forbes RB, Colville S, Swingle RJ: Cause of death and clinical grading criteria in a cohort of amyotrophic lateral sclerosis cases undergoing autopsy from the Scottish motor neurone disease register. *J Neurol Neurosurg Psychiatry* 2009;80:84–87.
- Wroe R, Wai-Ling Butler A, Andersen PM, Powell JF, Al-Chalabi A: ALSoD: the amyotrophic lateral sclerosis online database. *Amyotroph Lateral Scler* 2008;9:249–250.
- Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, Donaldson D, Goto J, O'Regan JP, Deng HX, et al: Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* 1993;362:59–62.
- Andersen PM, Restagno G, Stewart HG, Chio A: Disease penetrance in amyotrophic lateral sclerosis associated with mutations in the SOD1 gene. *Ann Neurol* 2004;55:298–299; author reply 299.
- Andersen PM, Sims KB, Xin WW, Kiely R, O'Neill G, Ravits J, Pioro E, Harati Y, Brower RD, Levine JS, Heinicke HU, Seltzer W, Boss M, Brown RH Jr: Sixteen novel mutations in the Cu/Zn superoxide dismutase gene in amyotrophic lateral sclerosis: a decade of discoveries, defects and disputes. *Amyotroph Lateral Scler Other Motor Neuron Disord* 2003;4:62–73.
- Shatunov A, Mok K, Newhouse S, Weale ME, Smith B, Vance C, Johnson L, Veldink JH, van Es MA, van den Berg LH, Robberecht W, Van Damme P, Hardiman O, Farmer AE, Lewis CM, Butler AW, Abel O, Andersen PM, Fogh I, Silani V, Chio A, Traynor BJ, Melki J, Meininger V, Landers JE, McGuffin P, Glass JD, Pall H, Leigh PN, Hardy J, Brown RH Jr, Powell JF, Orrell RW, Morrison KE, Shaw PJ, Shaw CE, Al-Chalabi A: Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *Lancet Neurol* 2010;9:986–994.
- Vance C, Al-Chalabi A, Ruddy D, Smith BN, Hu X, Sreedharan J, Siddique T, Schelhaas HJ, Kusters B, Troost D, Baas F, de Jong V, Shaw CE: Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2–21.3. *Brain* 2006;129:868–876.
- Morita M, Al-Chalabi A, Andersen PM, Hosler B, Sapp P, Englund E, Mitchell JE, Habgood JJ, de Bellerocche J, Xi J, Jongjaroenprasert W, Horvitz HR, Gunnarsson LG, Brown RH Jr: A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia. *Neurology* 2006;66:839–844.
- van Es MA, Veldink JH, Saris CG, Blauw HM, van Vught PW, Birve A, Lemmens R, Schelhaas HJ, Groen EJ, Huisman MH, van der Kooij AJ, de Visser M, Dahlberg C, Estrada K, Rivadeneira F, Hofman A, Zwarts MJ, van Doormaal PT, Rujescu D, Strengman E, Giegling I, Muglia P, Tomik B, Slowik A, Uitterlinden AG, Hendrich C, Waibel S, Meyer T, Ludolph AC, Glass JD, Purcell S, Cichon S, Nothen MM, Wichmann HE, Schreiber S, Vermeulen SH, Kiemeny LA, Wokke JH, Cronin S, McLaughlin RL, Hardiman O, Fumoto K, Pasterkamp RJ, Meininger V, Melki J, Leigh PN, Shaw CE, Landers JE, Al-Chalabi A, Brown RH Jr, Robberecht W, Andersen PM, Ophoff RA, van den Berg LH: Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat Genet* 2009;41:1083–1087.
- Van Deerlin VM, Sleiman PM, Martinez-Lage M, Chen-Plotkin A, Wang LS, Graff-Radford NR, Dickson DW, Rademakers R, Boeve BF, Grossman M, Arnold SE, Mann DM, Pickering-Brown SM, Seelaar H, Heutink P, van Swieten JC, Murrell JR, Ghetti B, Spina S, Grafman J, Hodges J, Spillantini MG, Gilman S, Lieberman AP, Kaye JA, Woltjer RL, Bigio EH, Mesulam M, Al-Sarraj S, Troakes C, Rosenberg RN, White CL, 3rd, Ferrer I, Llado A, Neumann M, Kretzschmar HA, Hulette CM, Welsh-Bohmer KA, Miller BL, Alzualde A, Lopez de Munain A, McKee AC, Gearing M, Levey AI, Lah JJ, Hardy J, Rohrer JD, Lashley T, Mackenzie IR, Feldman HH, Hamilton RL, Dekosky ST, van der Zee J, Kumar-Singh S, Van Broeckhoven C, Mayeux R, Vonsattel JP, Troncoso JC, Krill JJ, Kwok JB, Halliday GM, Bird TD, Ince PG, Shaw PJ, Cairns NJ, Morris JC, McLean CA, DeCarli C, Ellis WG, Freeman SH, Frosch MP, Growdon JH, Perl DP, Sano M, Bennett DA, Schneider JA, Beach TG, Reiman EM, Woodruff BK, Cummings J, Vinters HV, Miller CA, Chui HC, Alafuzoff I, Hartikainen P, Seilhean D, Galasko D, Masliah E, Cotman CW, Tunon MT, Martinez MC, Munoz DG, Carroll SL, Marson D, Riederer PF, Bogdanovic N, Schellenberg GD, Hakonarson H, Trojanowski JQ, Lee VM: Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat Genet* 2010;42:234–239.
- Boxer AL, Mackenzie IR, Boeve BF, Baker M, Seeley WW, Crook R, Feldman H, Hsiung GY, Rutherford N, Laluz V, Whitwell J, Foti D, McDade E, Molano J, Karydas A, Wojtas A, Goldman J, Mirsky J, Sengdy P, Dearmond S, Miller BL, Rademakers R: Clinical, neuroimaging and neuropathological features of a new chromosome 9p-linked FTD-ALS family. *J Neurol Neurosurg Psychiatry* 2011;82:196–203.
- Laaksovirta H, Puralinna T, Schymick JC, Scholz SW, Lai SL, Myllykangas L, Sulkava R, Jansson L, Hernandez DG, Gibbs JR, Nalls MA, Heckerman D, Tienari PJ, Traynor BJ: Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. *Lancet Neurol* 2010;9:978–985.

- 23 Le Ber I, Camuzat A, Berger E, Hannequin D, Laquerriere A, Golfier V, Seilhean D, Viennet G, Couratier P, Verpillat P, Heath S, Camu W, Martinaud O, Lacomblez L, Vercelletto M, Salachas F, Sellal F, Didic M, Thomas-Anterion C, Puel M, Michel BF, Besse C, Duyckaerts C, Meininger V, Campion D, Dubois B, Brice A: Chromosome 9p-linked families with frontotemporal dementia associated with motor neuron disease. *Neurology* 2009;72:1669–1676.
- 24 Pearson JP, Williams NM, Majounie E, Waite A, Stott J, Newsway V, Murray A, Hernandez D, Guerreiro R, Singleton AB, Neal J, Morris HR: Familial frontotemporal dementia with amyotrophic lateral sclerosis and a shared haplotype on chromosome 9p. *J Neurol* 2011;258:647–655.
- 25 Luty AA, Kwok JB, Thompson EM, Blumbergs P, Brooks WS, Loy CT, Dobson-Stone C, Panegyres PK, Hecker J, Nicholson GA, Halliday GM, Schofield PR: Pedigree with frontotemporal lobar degeneration-motor neuron disease and Tar DNA binding protein-43 positive neuropathology: genetic linkage to chromosome 9. *BMC Neurol* 2008;8:32.
- 26 Jefferies J: Births in England and Wales – Cohort Fertility 2009, Statistical Bulletin. Office of National Statistics, 2010.
- 27 Kane P, Choi CY: China's one child family policy. *BMJ* 1999;319:992–994.
- 28 Ginsburg E, Malkin I, Elston RC: Sampling correction in pedigree analysis. *Stat Appl Genet Mol Biol* 2003;2: Article 2.
- 29 Burton PR, Palmer LJ, Jacobs K, Keen KJ, Olson JM, Elston RC: Ascertainment adjustment: where does it take us? *Am J Hum Genet* 2000;67:1505–1514.
- 30 So HC, Gui AH, Cherny SS, Sham PC: Evaluating the heritability explained by known susceptibility variants: a survey of ten complex diseases. *Genet Epidemiol* 2011;35:310–317.
- 31 Lee SH, Wray NR, Goddard ME, Visscher PM: Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 2011;88:294–305.