Ellis-van Creveld syndrome and the Amish

Victor A. McKusick

Institute of Genetic Medicine, Johns Hopkins Hospital, Baltimore, Maryland 21287, USA. e-mail: mckusick@peas.welch.jhu.edu

Genetic studies often involve the cooperation of large numbers of affected persons and their families. The discovery of the gene that, when mutated, causes a form of dwarfism (Ellis-van Creveld syndrome) has been accelerated through a collaborative effort between geneticists and the Old Order Amish, of Lancaster County, Pennsylvania.

As described on pages 283–286 of this issue, Judith Goodship and a multinational group of collaborators[1] have identified the gene that is mutated in people with a form of dwarfism, Ellis-van Creveld (EvC) syndrome. They discovered five different mutations, including one that underlies the disorder in the Old Order Amish. Originally described in 1940 (ref. 2) by paediatricians Richard Ellis and Simon Van Creveld, EvC syndrome is an autosomal recessive disorder, involving postaxial polydactyly of the hands (see figure), short stature with shortening especially of the forearms and lower legs and, in at least half of all cases, congenital heart malformation.

The mutation in the Amish of Lancaster County, Pennsylvania, in whom the disorder occurs at unprecedentedly high frequency, is predicted to cause aberrant splicing. It occurs in the fifth nucleotide of intron 13 of a novel gene, *EVC,* that is predicted to encode a protein containing a leucine zipper, three putative nuclear localization signals and a putative transmembrane domain. The pathogenic 'status' of the Amish mutation is supported by the fact that mutation of analogous nucleotides effect disease: according to the Human Gene Mutation Database[3], 144 intronic mutations, causing a total of 81 separate disorders, have been reported at the +5 position of introns.

Amish mother and child. The child has Ellis-van Creveld syndrome, which is characterized by polydactyly (six fingers on each hand), short stature, and shortening of the forearms and lower legs. (Image reproduced with permission from Johns Hopkins University Press).

Of history and heritage

The Amish have several characteristics[4,5] that recommend them to medical geneticists. First, they are descended from a limited number of founders who immigrated, during the eighteenth century, to the United States from the Rhineland (in the southwest of Germany) where they had settled temporarily following emigration in the 1690s from the Canton of Berne, Switzerland. Second, the Amish observe strict endogamy (they marry only within the community), with gene flow being exclusively centrifugal (that is, members may leave the community but 'outsiders' do not join it and thereby introduce exogenous genes). Third, like the Icelanders, they keep excellent genealogic records and have a restricted geography. Finally, they tend to have large families, with many children.

It was therefore possible to trace the lineage of both parents of all 50 EvC cases back to a single couple, Samuel King and his wife (regrettably, her name is no longer known), who immigrated[4] to Eastern Pennsylvania in 1744—thus demonstrating founder effect and a recessive pattern of inheritance. Epidemiological data indicate that the frequency of the mutated gene is approximately 0.066 and that heterozygotes make up about 12.3% of the population[6]. At the time of these estimates, 12.6% of Lancaster County Amish carried Samuel King's surname, and Samuel was the only male founder of that name.

During studies carried out in the mid-sixties[5,7], it became apparent that the Amish are distributed in three consanguineal kin groups[8] (demes) across the United States. At that time, each was made up of about 14,000 members. The deme in Lancaster County was founded by those who immigrated before the American Revolution. The deme of Holmes County (Ohio), and the deme comprised of groups in Lagrange and Elkhart Counties (Northern Indiana) descended, for the most part, from post-revolutionary immigrants who, upon finding the land taken up in Eastern Pennsylvania, moved to points west.

The genetic distinctness of the three major demes is supported by different patterns of blood-group frequencies[9,10], different family names—23% of people in the Lancaster deme have the name Stoltzfus, which is absent in the other demes—and different frequencies of rare recessive disorders. For example, EvC syndrome was found to be limited to the Lancaster-County deme. Haemophilia B, on the other hand, was (and still is) unusually frequent in the Holmes-County deme, and almost completely limited to that group. It is as though the Amish immigrants were streaked like bacteria on a culture plate across the waist of America, with the genetic profile of each deme depending on the genetic constitution of the founders, for whom the present populations represent a bioassay.

Cartilage-hair hypoplasia

A second recessive form of dwarfism, distinct from EvC, is prevalent in the Lancaster-County deme[11]. Cartilage-hair hypoplasia (CHH) was previously unrecognized until the Amish came to the attention of clinical geneticists in the mid-sixties. In contrast with EvC syndrome, it occurs in all Amish demes. Moreover, it is impossible to trace its origin to a single founder couple, indicating that the mutation was introduced by several immigrants. It turns out that CHH is also frequent in Finland[12]; the odds favour a mutation of independent origin with

respect to that carried by the Amish—albeit one that has achieved a high frequency through the same mechanisms: founder effect and perhaps genetic drift. Whereas the 'causative' gene's locus is known[13], its identity yet eludes the assiduous efforts of positional cloners in Helsinki and Bethesda.

After discovery of CHH in the Amish, rare cases of CHH were recognized in non-Amish. For example, Billy Barty, an actor and founder of Little People of America, a support group for persons of short stature, has CHH. So did Michael ('Pat') Bilon, who played ET in the movie of that name.

Medical genetics is indebted to the Amish for their cooperation in studies that have led to an improved understanding of genetic disorders. The physicians who carried out the studies in the 1960s and 70s approached the Amish with a view to helping them. Arrangements were made, for example, for surgical repair of the cardiac defect in EvC patients and for orthopedic correction of their knee deformities. Aid was also provided to family members with non-EvC related problems of great diversity. How could knowledge of the Amish 'EvC' mutation help? Premarital and pre-natal counselling should now be possible, based on testing for the splice-site mutation or a nearby marker—ideally one within the gene. Goodship and colleagues discovered a polymorphism that is in linkage disequilibrium with the 'causative' mutation, and could therefore serve as such a marker.

Whether the Amish would acquiesce to premarital testing is uncertain, and it is unlikely that they would accept prenatal testing because of the implication of abortion. Because a specific *EVC* mutation is limited to the Lancaster County Amish, marriage between an EvC carrier with an Amish from another community might be recommended but may generate logistical difficulties. Alternatively, knowledge of carrier status could inform choice of partner within the Lancaster County Amish community.

The Amish acceptance of the geneticists was achieved by their being introduced by local physicians and by sociologists whom they trusted. The relationships were maintained through communication with the bishops and others in authority and by the assistance of Amish who served as guides and introducers during home visits. Another notable example of beneficial collaboration between geneticists and religous community is that between the Ashkenazi Jewish groups who use screening for mutations that cause Tay-Sachs disease as the basis of marriage advice by rabbis.

The EvC syndrome in the Amish has become a favourite elementary genetics textbook example of several aspects of human genetics. Now, to founder effect, consanguinity, recessive inheritance and so on, one can add linkage mapping, positional cloning and the molecular nature of mutation, as well as carrier detection and the social implications thereof. Possibly, it will not be long before the student can be informed of the way in which the mutation disturbs development, leading to polydactyly, heart defect and skeletal dysplasia. □

1. Ruiz-Perez, R.-L. *et al. Nature Genet.* **24**, 283–286 (2000).
2. Ellis, R.W.B. & van Creveld, S. *Arch. Dis. Child.* **15**, 65–84 (1940).
3. Krawczak, M. & Cooper, D.M. *Trends Genet.* **13**, 121–122 (1997).
4. Hostetler, J.A. *Amish Society* (Johns Hopkins University Press, Baltimore, 1993).
5. McKusick, V.A., Hostetler, J.A., Egeland, J.A. & Eldridge, R. *Cold Spring Harbor Symp. Quant. Biol.* **29**, 99–113 (1964).
6. McKusick, V.A. in *Medical Genetic Studies of the Amish: Selected Papers Assembled, with Commentary* (ed. McKusick, V.A.) 135–139 (Johns Hopkins University Press, Baltimore, 1978).
7. Cross, H.E. & McKusick, V.A. *Social Biol.* **17**, 83–101 (1970).
8. Murdock, G.P. *Social Structure* (Macmillan, New York, 1949).
9. McKusick, V.A., Bias, W.B., Norum, R.A. & Cross, H.E. *Humangenetik* **5**, 36–41 (1967).
10. Juberg, R.C., Schull, W.J., Gershowitz, H. & Davis, L.M. *Human Biol.* **43**, 477–485 (1971).
11. McKusick, V.A., Eldridge, R., Hostetler, J.A., Ruangwit, U. & Egeland, J.A. *Bull. Johns Hopkins Hosp.* **116**, 285–286 (1965).
12. Kaitila, I. & Perheentupa, J. in *Population Structure and Genetic Disorders* (eds Eriksson, A.W., Forsius, H.R., Nevanlinna, H.R., Workman, P.L. & Norio, R.K.) 588–591 (Academic, New York, 1980).
13. Sulisalo, T. *et al. Nature Genet.* **3**, 338–341 (1993).

Making the most of microarray data

Terry Gaasterland & Stefan Bekiranov

Laboratory of Computational Genomics, The Rockefeller University, 1230 York Ave, New York, New York, 10021, USA.
e-mail: gaasterland@genomes.rockefeller.edu and bek@genomes.rockefeller.edu

The impact of microarray technology on biology will depend on computational methods of data analysis. A supervised computer-learning method using support vector machines predicts gene function from expression data—and shows promise.

Microarray assays can measure the transcriptional effects of changes in gene function under different conditions. They can reveal genes that characterize tissue type, developmental stage, or responses to environmental conditions or genetic modifications. Microarray assays will therefore become a general feature of experimental protocols in genetics and cell physiology. As array data burgeon, new questions arise: if we, as a research community, collect all array hybridization data in a central location[1], can we assign new genes of unknown function to known functional classes? Can we correlate gene expression with gene function? Can we find new classes of co-regulated genes? Can we extract complete gene regulatory networks from microarray gene expression data?

Computation is our only hope, and an article by Michael Brown and colleagues[2] in a recent issue of *The Proceedings of the National Academy of Sciences* describes an approach to microarray data analysis that addresses the first question. The authors use support vector machines (SVMs; Fig. 1), a supervised computer-learning method, to train a 'classification machine' to recognize new genes that are similar in expression pattern to groups of genes known to be co-regulated. In contrast with classical unsupervised clustering methods and pure self-organizing maps, the approach builds on existing knowledge (Fig. 2) and has the potential to refine and correct it.