The phenotype/genotype relation and the current status of genetic screening in hypertrophic cardiomyopathy, Marfan syndrome, and the long QT syndrome

J Burn, J Camm, M J Davies, L Peltonen, P J Schwartz, H Watkins

Hypertrophic cardiomyopathy, Marfan syndrome, and the long QT syndrome are all autosomal disorders inherited in a dominant manner; affected individuals are heterozygous, that is they have one normal and one mutant copy of the gene but not all the gene carriers are symptomatic. This failure to express fully the expected phenotype is traditionally known as incomplete penetrance.

Vocabulary and technology

In the human genome (consisting of 22 pairs of autosomes and the X and Y chromosomes) there are between 50 000 and 100 000 individual genes coding for a wide range of products including enzymes and structural proteins. Each chromosome is paired, giving a maternal and a paternal contribution to each gene. Each gene consists of a series of base pairs, the sequence of which is used as the template for producing the messenger RNA, which is used in turn to synthesise proteins on the ribosomes in the cytoplasm. There are four bases: adenine (A), guanine (G), cytosine (C), and thymine (T). A occurs opposite T and C opposite G in the two strands of the DNA double helix; one, the sense strand, is transcribed; the other (antisense) is not. The sequences in each strand are complementary but run in opposite directions. The two copies (alleles) of the gene in an individual from each parent consist of a segment of the double stranded DNA molecules. The base sequence in the DNA strand is a triplet code in which sets of three bases determine an amino acid. In the messenger RNA molecule, thymine (T) is replaced with uridine (U). Of the 64 possible coding triplets which can be formed from the four bases, some define gene boundaries: TAG and TGA are termination signals, and AUG an initiation signal. The other triplets code for a specific amino acid, but there is considerable redundancy; there are more triplet permutations than amino acids.

Genes vary widely in size from a few hundred to many thousand base pairs. The genes are widely spaced and between them are base sequences with no known function. Within each gene there are sequences which direct gene expression, and sequences (exons) which encode amino acids, but within the gene there are also non-functional sequences (introns).

The simplest single gene abnormality is a point mutation in which one base pair is replaced by another. Because of the redundancy in the genetic code, point mutations may be silent polymorphisms that do not change the amino acid or they may be missense mutations that change a single amino acid. Missense mutations can be regarded as analogous to simple spelling errors in one letter of a word. For example “the fish is on the dish” becomes “the fish is on the fish”. Because each gene exists in a maternal and a paternal copy the ultimate product of the gene is a combination of normal (wild type) and abnormal (mutant type). If the subject has inherited an abnormal gene from both mother and father (homozygous) the product is entirely of the mutant type. Some of the earliest recognised genetic disorders, such as sickle cell anaemia, followed this simple inheritance pattern, carriers (heterozygotes) having a mixture of adult (A) and sickle (S) haemoglobin, homozygotes only have haemoglobin S.

Other abnormalities of the gene structure include deletions of one or more base pairs. Such deletions can have a major effect if anything other than a multiple of three base pairs is removed because the triplet will be read out of frame, which will scramble the code thereafter. By analogy, loss of one letter from the text sequence used above while the word length remained the same would make “Thf ishi so nt bed isp”.

Where an abnormally situated stop codon results from a point mutation or from a frame shift translation of the mRNA product of this gene, a truncated protein product or no product is synthesised. More recently yet another class of mutation has been identified: triplet repeat sequences may be abnormally reduplicated, with marked increase in length of the gene until it is ultimately inactivated. Such repetition may progress from generation to generation so that the disease phenotype gets worse (anticipation). An example is myotonic dystrophy.

Individuals who are homozygous will have severe disease because either no gene product is produced or the product is all of the mutant...
type. Homozygotes require both parents to be abnormal gene carriers. This is rare unless the parents are related or the frequency of the mutant gene is very high in the general population. With mutations of structural protein genes, disease may be manifest in heterozygotes because the presence of mutant product interferes with the function of the wild type product (termed a dominant negative action).

Tools in gene identification
A widely used technique for the initial identification of a previously unknown mutated disease related gene is linkage analysis. When large families are available in which there are two or more generations having clinically affected and non-affected individuals, study of DNA extracted from peripheral blood cells can lead to identification of the chromosome on which the gene is situated. This is based on the utilisation of the polymorphic markers in the human genome. Over 5000 polymorphic non-functional DNA sequences (markers) are evenly spread on individual chromosomes and form the basis of the first generation map of the human genome. A known polymorphic marker which systematically segregates with the disease in the family flags the segment of a chromosome where the disease gene lies. The human genome project provides rapidly increasing amounts of data on genes residing in that area. However, there are often many of these; the products which such genes code for can, however, be considered and a guess made about the most likely candidate gene for a particular clinical picture. In the final stage DNA sequences are identified in which there is a change in base sequence which is always inherited with the disease, is not present in unaffected individuals, and has an abnormal protein product which provides a reasonable explanation for the phenotype.

Linkage analysis, while a powerful tool, depends on having large families containing both affected and normal individuals. Several factors may disturb what is ostensibly a logical approach that is certain of success. In the first place, very accurate clinical establishment of whether any individual is or is not carrying the gene is vital. Linkage becomes difficult if some family members carry the gene but have no clinical manifestations. Second, key family members may have died or moved to other countries, so that blood samples are not available; a final problem is that up to 5% of individuals in genetic studies can be shown not to have the father named on their birth certificate.

Genotype-phenotype relations
The understanding of genotype-phenotype relations is proving more complex than originally envisaged. A three dimensional approach is needed. One aspect is the clinical picture, that is, the phenotype. This involves the structural changes observed and the nature and severity of the clinical disease. The second aspect is the gene itself; its DNA sequence, and what protein it codes. The last aspect involves an understanding of how the abnormal protein product produces the phenotype. The whole picture is not yet clear for hypertrophic cardiomyopathy, Marfan disease, or the long QT syndrome, but major advances have been made and the three dimensional construct of phenotype, gene, and product function is proving somewhat different for each.

Hypertrophic cardiomyopathy
Point mutations in five separate genes are currently known to produce hypertrophic cardiomyopathy. The phenotype is characterised by left ventricular wall thickening with characteristic histological appearance known as myocyte disarray, which separates the condition from left ventricular hypertrophy secondary to hypertension or aortic stenosis. Disarray occurs at two levels—within the cell, myofibrillar structure is disorganised with crossing rather than the normal parallel array; the myocytes are also arranged abnormally in relation to each other, running in whorls around foci of connective tissue. In areas with disarray the myocytes are misshapen with incomplete aberrant cell junctions occurring over much of the surface rather than being confined to the intercalated discs at the ends of a roughly oblong shaped cell.

The current model of hypertrophic cardiomyopathy is that the known genes code for proteins concerned with the myofibril itself. These are cardiac myosin heavy chain (chromosome 14), cardiac troponin T (chromosome 1), troponymosin (chromosome 15), and myosin binding protein C (chromosome 11). The protein coded for by a gene on chromosome 7 is unknown but a phenotype indistinguishable from the others is produced. The gene on chromosome 7 is tightly linked to pre-excitation in a family. This implies either that the gene produces both hypertrophic cardiomyopathy and pre-excitation, or that there are two immediately adjacent genes which are rarely separated in meiosis.

Virtually all cases of hypertrophic cardiomyopathy are heterozygous and the gene is a dominant negative. This means that the mutant myofibrillary protein interferes with the function of the wild (normal) protein in forming properly aligned myofibrils within the myocyte. In reports on hypertrophic cardiomyopathy worldwide, there are few examples of proven new mutations in which neither parent carried the abnormal gene. This means that in the majority of families with an apparent sporadic case of hypertrophic cardiomyopathy there are likely to be other asymptomatic gene carriers. In the past, series which identified carriers by the clinical expression alone failed to detect many gene carriers. β Heavy chain myosin mutations account for around 35% of cases of hypertrophic cardiomyopathy and over 40 separate point mutations in the gene are now known. Mutations in the other genes are less common, each accounting for 5–15% of cases. At least eight mutations are now described in cardiac troponin T, three in tropomyosin, and three, including two splice site mutations that lead to a frame shift in the gene, for myosin binding protein C.
The challenge in the full understanding of the genotype/phenotype relation in hypertrophic cardiomyopathy is the wide range of the phenotype and clinical effects seen in individual subjects. The earliest accounts of the disease recognised only the phenotype associated with asymmetric septal thickening and left ventricular outflow obstruction. It is now apparent that the phenotype is very variable. The hypertrophy may be symmetric or asymmetric, there may or may not be concomitant right ventricular involvement, and the degree of increase in left ventricular mass varies widely. Some cases have very minimal left ventricular wall thickening.

In some degree the variation in phenotypic expression and clinical outcome is dependent on the specific gene and the mutation. In the heavy chain myosin gene some specific mutations carry a very high risk of sudden death and have a “malignant” family history. Troponin T mutations produce far less striking left ventricular hypertrophy than the average heavy chain myosin mutation but carry a high risk of sudden death.15 Many families with hypertrophic cardiomyopathy, however, show the full range of phenotypic appearances even though all the members carry precisely the same mutation. The cause for this intrafamily variation is unknown. The ratio of wild to mutant protein within the myocardiurn, or whether it varies from case to case within a family, is unknown. The reason why some cases show striking regional ventricular involvement with a histologically abnormal septum and a normal posterior wall is unknown. Local haemodynamic factors might account for the preponderance of septal thickening but the most likely explanation for interindividual variation is that other genes impinge on the phenotypic expression of the hypertrophic cardiomyopathy gene. Any gene that enhanced hypertrophy would be expected to potentiate the phenotypic expression of hypertrophic cardiomyopathy.

There is considerable selection bias in published reports concerning hypertrophic cardiomyopathy. Families referred to tertiary centres for genetic investigation usually represent the severe cases where overt disease is present in many members. The prognosis and severity of disease in cases seen outside specialist tertiary referral centres may be better.

The potential pathways along which any subject carrying one of the hypertrophic cardiomyopathy genes can progress are well known (figure) but what determines which patient and which path is unknown. Sudden death is always a risk at any age, irrespective of the presence or absence of symptoms. For example a study of 150 cases presenting with minimal symptoms showed an overall frequency of sudden death of 1-5% per annum.7 Another similar study of 202 cases of hypertrophic cardiomyopathy8 put the figure as low as 0-6% per annum. Nevertheless the rate of sudden death is higher in many series from tertiary referral centres. Factors indicating higher risk are a history of two or more deaths in the family, early onset of symptoms, a fall of blood pressure on exercise, and severe left ventricular wall thickening (> 35 mm). Synchronic attacks due to non-sustained ventricular tachycardia also indicate high risk, and in this subgroup amiodorone may reduce the risk.9

**Marfan syndrome**

Like hypertrophic cardiomyopathy, Marfan syndrome is an autosomal dominant with a very wide phenotypic range both within affected families and between families.10 Two genes encode for components of fibrillin—the common mutations producing Marfan syndrome involve the fibrillin gene on chromosome 15.11-14 Fibrillin is a 350 kDa protein produced by connective tissue cells and secreted into the extracellular matrix, where it forms fibrillar aggregates which serve in adhesion of connective tissue structures. The gene is huge, being comprised of many thousands of base pairs encoding numerous repetitive motifs.14,15 Common repeats include an epidermal growth factor (EGF) like and transforming growth factor β (TGF β) binding protein like motifs. The phenotypic range is very wide—abnormal fibrillin is responsible for all the phenotypes listed in the table. Mutations in exons 24–32, containing the longest continuous stretch of EGF repeats, are particularly liable to cause a severe defect, often fatal in the neonatal period.16

Congenital contractual arachnodactyly is due to mutations in the second fibrillin gene on chromosome 5.17 A phenotype very close to Marfan syndrome has also been reported in single families due to a mutation in the gene coding for α2(1) collagen and an unknown
gene on chromosome 3. Some families breed true, that is, have all members confined to one phenotypic group, while other families show wide variation.\textsuperscript{14,15} Cardiac involvement can occur in all the subtypes.

In large part the severity of the phenotype depends on the alterations in the amount of wild type versus mutant fibrillin that is produced. Some mutations have a classic dominant negative effect in that mutant fibrillin interferes with the function of the normal protein. This can be tested in vitro—fibroblasts grown from skin in a skin biopsy produce a regular network of microtubules that can be seen as bundles by electron microscopy.\textsuperscript{16,19} Fibroblasts from subjects carrying abnormal fibrillin genes produce microfibris which form thinner fragmented strands—very severe cases produce no microfibris. The phenotype thus depends on the ratio of wild type to mutant fibrillin which in turn depends on the nature of the mutation. Over 80 single point mutations deletions and insertions are known—so far every Marfan family has proved unique, a reflection of the huge size of the gene. Relatively severe disease seems to occur with proportions of mutant fibrillin exceeding the threshold of approximately 10%. Many missense deletions, premature stop codons, or whole segment deletions produce either no mutant fibrillin or relatively small proportions of mutant type, and the clinical phenotype is then mild. For example in the mitral aortic skeletal skin (MASS) syndrome the patients lack the full clinical criteria for Marfan syndrome but have an insertion into the fibrillin gene leading to a premature stop codon. The level of mutant fibrillin is as low as 6% and symptoms mild.

In Marfan syndrome, therefore, while many cases are dominant negatives with mutant fibrillin interfering with normal fibrillin, others are null alleles, that is, there is a simple deficiency of fibrillin.

The long QT syndrome

At present the most straightforward genotype/phenotype link occurs in the long QT syndrome.\textsuperscript{20} Even here, however, the phenotype is broader than initially envisaged. The proband in families usually presents with syncope due to ventricular tachyarrhythmias or cardiac arrest on exercise or emotional stress, often under the age of 15. Within the family asymptomatic gene carriers may also have a long QT interval; of these up to 20% ultimately become symptomatic. In some families, however, a few carriers may have a normal QT interval—these individuals have a lower but still definite risk of syncope. The mortality in untreated symptomatic cases exceeds 60% in 15 years. Effective treatments are already available and mortality has been reduced to approximately 3–4% within 10 years from the first episode. The mainstays of treatment are \( \beta \) blockers, left cardiac sympathetic denervation,\textsuperscript{21} and pacing whenever there is clear evidence of pause induced or bradycardia dependent syncope.\textsuperscript{18} In subjects who die, the heart is totally normal by all morphological and structural criteria. In such cases the cause of death at necropsy, unless there is prior ECG data, cannot be ascertained. In these cases screening by ECG of surviving relatives is the only way to establish the diagnosis.

Linkage has been established on four chromosomes: 11, 7, 3, and 4\textsuperscript{22–24} but a significant number of families appear linked to none of these chromosomes.\textsuperscript{25} Between 1995 and 1996 the genes for LQTI (long QT syndrome linked to chromosome 11), for LQT2 (long QT syndrome linked to chromosome 7), and for LQT3 (long QT syndrome linked to chromosome 3) have been identified.\textsuperscript{26–29} The gene for LQT is KVLTQ, which, when coassembled with min K\(_e\) encodes for the I\(_a\) current.\textsuperscript{26,29} The gene for LQT2 is HERO, a potassium channel that carries the I\(_a\) current. The gene for LQT3 is SCNSA, the cardiac sodium channel gene and the three mutations described so far\textsuperscript{30} affect a region thought to be important for sodium inactivation. It is tempting to regard LQT2 and LQT3 as mutations in the same gene; however, the fifth locus for the long QT syndrome is a mutation in the min K gene on chromosome 21.

The mutated genes for LQT2 and LQT3 have been expressed in Xenopus oocytes\textsuperscript{32–34}, this has represented a major step forward because it has allowed the first demonstration of an alteration in function due to a specific mutation.

Expression of the mutant SCNSA genes has shown that they result in a small sustained inward current which is likely to be sufficient to disrupt the normal balance between inward and outward currents during the plateau phase, and hence prolong cardiac action potentials. It has practical importance in that this persistent inward current is blocked by mexiletine, a sodium channel blocker.\textsuperscript{34} Similarly, in a cellular model mimicking the SCNSA and HERG defects present in LQTS, it has been observed that mexiletine corrects the action potential prolongation induced by anthopleurin, a blocker of sodium channel inactivation used to mimic LQT3, but that not induced by dofetilide, an I\(_a\) blocker used to mimic LQT2.\textsuperscript{35}

Expression of the mutant HERG genes has led to the identification of two main consequences.\textsuperscript{36} Some of the mutant proteins do not form functional channels and do interact with normal HERG channel when expressed in Xenopus oocytes. This implies that patients with these mutations will probably express half the normal number of channels carrying I\(_a\). Other mutant channels do not express detectable currents but cause a dominant negative suppression of the normal HERG function; thus patients with these mutations will have a major reduction in I\(_a\) with a large effect on ventricular repolarisation.

The expression studies that have succeeded in comparing different mutations\textsuperscript{33,34} indicate that the magnitude of the effect on repolarisation varies according to the specific mutation. It will be interesting to determine if there is a correlate in the severity of the clinical manifestations and the spectrum of either HERG or...
SCNSA dysfunction in patients with LQT2 and LQT3. Patients with LQT1, LQT2, and LQT3 have a different electrocardiographic phenotype with a different shape of the T wave. Attempts are being made to correlate the various mutations with the clinical responses. The sodium channel blocker mexiletine produced a considerable shortening of the QT interval in LQT3 patients, but not in LQT2 and LQT1 patients. Even though a QT shortening is by no means a guarantee of protection from life-threatening arrhythmias, these data suggest that it may be appropriate to test the potential value of mexiletine specifically in the patients with the SCNSA mutations.

Heart rate increase produced a rather marked shortening of the QT interval among LQT3 patients. This effect was much less evident among LQT2 and LQT1 patients. The inference could be that LQT3 patients may be at lesser risk of syncope during physical exercise, when the progressive heart rate increase may allow appropriate QT shortening. These patients may also be those less likely to be protected by \( \beta \) blockers, which would maintain an excessively low heart rate.

Finally, in these initial studies it has been possible to observe an association between physical or emotional stress and syncpe for LQT1 patients, whereas LQT3 patients seem to be more at risk during rest or sleep. LQT2 patients are between these two extremes.

These observations should be accepted with caution because the small number of patients studied so far does not allow extrapolation to the entire population affected by long QT syndrome and overlap certainly exists. Nonetheless, this study has provided the first demonstration of differential responses of patients with long QT syndrome to interventions targeted to their specific genetic defect.

The long QT syndrome represents a model of correlation between specific mutations and specific alterations in function, thus representing the best example available so far of genotype-phenotype correlations. Further development is likely to have an impact on the management of these patients by allowing novel and highly specific interventions to be tested.

**Genetic screening—when is it indicated?**

The cornerstone of managing a patient with hypertrophic cardiomyopathy, Marfan disease, or the long QT syndrome is a full family history of the ages and causes of death. These histories are a major indicator for assessing the risk of sudden death in the individual.

The taking of such histories will, however, inevitably alert the patient to the fact the disease is familial. If there have been sudden deaths or aortic dissection at a young age in the family there is likely to be anxiety and requests for advice on who else is at risk, the magnitude of the risk, and whether anything can be done. Echocardiography and electrocardiography will detect only a proportion of gene carriers in all three conditions. In children in particular, phenotypic expression detectable by echocardiography may be absent and only appear later in life.

Genetic screening is at the moment expensive and its purpose must always be questioned. One purpose is clear—the extension of scientific knowledge, which is likely in the future to benefit all individuals with the abnormal gene. Unknown genes can only be identified by linkage analysis within large families which have, therefore, become an extremely valuable research resource. Science has become competitive and there are pressures to study families so as to be the first to describe a new gene or mutation. The question is whether this knowledge is beneficial, neutral, or harmful to both the family as a whole or to individuals within it.

One school of thought discourages the screening of asymptomatic children to detect gene carriers. The advice is heavily biased by the question of gene carriage in diseases such as Huntington's chorea, in which the disease is not manifest until middle age and nothing can be done in terms of treatment. It is felt that parents do not have the right to decide for their children whether they wish to be tested. The three cardiological genes discussed here do not, however, fit well into this scenario, and gene testing has potential benefits.

The situation is perhaps clearest for the long QT syndrome. The different long QT genes differ in the phenotype they produce, which in turn may determine the class of drug to use. There is clear evidence that drug treatment will reduce the risk of sudden death and that asymptomatic gene carriers are also at risk. The benefit to individuals or families of knowing who has the gene are therefore well established.

The situation is also clear in Marfan disease. The gene for fibrillin is so large that every family is unique with regard to their own mutation. To sequence the gene in a Marfan family is a major research undertaking not a routine clinical tool. Marfan families have to be managed by simple clinical techniques such as echocardiography to detect developing aortic root dilatation. Care have to be detected by physical measurements of hands, feet, aortic root, and so on. There is a case to be made for centres developing specific clinical expertise in managing the cardiovascular complications of Marfan disease but not for a central genetic screening service at the moment. This situation may, however, change when automated methods of screening the whole fibrillin gene become established and widely available. In vitro testing of fibroblasts in culture shows abnormal fibrillin structure, amount, or metabolism. This can help establish the diagnosis in clinically ambiguous cases and, if carried out in children, give some indication of the expected severity of disease without having to identify the precise mutation involved.

Hypertrophic cardiomyopathy is currently the most uncertain with regard to the need to detect gene carriers. It is technically feasible to screen all of the known contractile protein disease genes to detect point mutations. With
current technology this is laborious; in practise such screening is largely limited to families with several affected individuals such that the relevant disease can be first identified by linkage analysis.

There is a risk of sudden death in asymptomatic gene carriers, even when echocardiographic abnormalities are absent or minor. The magnitude of risk of sudden death, even in asymptomatic gene carriers, is higher with some specific mutations of the heavy chain myosin gene than others and is also high with the troponin T gene. These are arguments for identifying the responsible gene and detecting carriers. Such high risk families will often, however, have already declared themselves by more than one sudden death. In symptomatic individuals who are already having syncopal attacks, investigation to identify the different mechanisms which produce syncope and treat accordingly is entirely justified. What is far less clear at the moment is whether the asymptomatic gene carrier has relatives of whom they can have their risk of sudden death reduced by either drug therapy or restriction of physical activity.

At a practical level this means that genetic study of the family of an apparently isolated living case of hypertrophic cardiomyopathy may not be justified. The management of a family in whom there is an unexpected sudden death under 50 years of age as the first manifestation of hypertrophic cardiomyopathy is not clearly defined. In the United Kingdom, many index cases will be seen by pathologists working for the coroner and their responsibilities need defining. If nothing is done and another sudden death occurs, the family will rightly feel aggrieved. It is therefore advisable to inform the family through their general practitioner, which will start the sequence of a detailed family history and lead almost inevitably to referral of siblings and offspring to a cardiologist for screening. It is unresolved at the moment whether such screening should go beyond echocardiography and ECG.

Ethical aspects of genetic screening

One deterrent to the development of genetic registers and predictive testing is the potential for eugenic abuse. In 1995 China passed a law prohibiting marriage if a couple carried genetic disorders. It is therefore mandatory that genetic services are developed for the benefit of individuals and their families rather than improving society or making savings in costs for the State.

The detection of gene carriers can cause considerable detriment to family members. Insurance cover may become difficult to obtain and companies may load premiums, even if the individual is shown not to be a carrier. The parent who has passed on the gene to children may develop feelings of guilt and may be made to feel guilty by the other parent. Siblings who do not carry the gene are spared any further medical surveillance and will usually be delighted. Guilt occurs over personal survival when a much loved sibling dies, however, or there may be anger if negative decisions over life style or having a family prove to have been needless. Those who find they do carry the gene may be converted to chronic invalidism or anxiety. It is reasonable to carry out a full genetic study when the family asks for it and unreasonable to deny it to them. This view is supported by a lay body (Genetic Interest Group)39 representing families. Two provisos prevail: screening must be linked to expert counselling, and family members who do not wish to take part must not be put under pressure.

Genetic screening—who do it do?

Genetic screening for hypertrophic cardiomyopathy and the long QT syndrome is at the moment largely carried out by tertiary referral centres in cardiology, with a specific research interest in molecular genetics. This research service will ultimately have to be taken over by the National Health Service in some form.

All three diseases discussed here have the potential to devastate families by sudden unexpected deaths, and families have the right to the best available advice. Few cardiologists have genetic expertise and equally few geneticists have the clinical experience either to manage cases or to identify carriers by clinical criteria. The future must lie in partnership and joint management. A recent confidential inquiry into counselling for genetic disorders30 has found that in all the non-cardiac clinical specialties looked at so far there is important understanding of, or action on, the requirements of patients with genetic problems. It is unlikely that cardiovascular related genetic disorders will prove to be different.

Supported by an educational grant from Bayer plc.

11 Pereira L, Levran O, Ramirez F, Lynch JR, Sykes B,


