Keratins are the intermediate filament proteins specifically expressed by epithelial cells. The Human Genome Project has uncovered a total of 54 functional keratin genes that are differentially expressed in specific epithelial structures of the body, many of which involve the epidermis and its appendages. Pachyonychia congenita (PC) is a group of autosomal dominant genodermatoses affecting the nails, thick skin and other ectodermal structures, according to specific sub-type. The major clinical variants of the disorder (PC-1 and PC-2) are known to be caused by dominant-negative mutations in one of four differentiation-specific keratins: K6a, K6b, K16, and K17. A total of 20 human keratin genes are currently linked to single-gene disorders or are predisposing factors in complex traits. In addition, a further six intermediate filament genes have been linked to other non-epithelial genetic disorders. We have established a comprehensive mutation database that catalogs all published independent occurrences of intermediate filament mutations (http://www.interfil.org), with details of phenotypes, published papers, patient support groups and other information. Here, we review the genotype–phenotype trends emerging from the spectrum of mutations in these genes and apply these correlations to make predictions about PC phenotypes based on the site of mutation and keratin pair involved.

Key words: age of onset/genodermatosis/keratoderma/nail dystrophy/steatocystoma

based on the date of acceptance of the relevant publication. Source references and abstracts are given for each reported mutation. Care is taken to avoid duplication of mutations that are mentioned in more than one paper (e.g., in a genetics paper and perhaps a related clinical case report), however, independent examples of the same mutations are included so that mutation clusters and hotspot codons can be recognized. The individual reference sequences used for mutation nomenclature are available within the database and these are reviewed frequently in light of revisions to the human genome sequence. Mutations and polymorphisms are named in accordance to the published HUGO guidelines (Antonarakis, 1998; Dunn et al., 2000) and where the authors have used non-standard nomenclature, this is corrected in the database entry with an annotation to show that there has been an alteration. Thus, the corrected data allows comparisons to be made without confusion. Furthermore, the database contains amino acid reference sequences that have been annotated to show the positions of functional protein domains. The protein domain affected is listed with individual mutations to allow comparisons to be made between different keratins or other intermediate filament genes. There are also brief clinical descriptions of the various intermediate filament disease phenotypes as well as links to the clinical genetics database OMIM, patient advocacy groups and other useful sites. It is hoped that this database will be a central resource for the field, allowing genotype–phenotype and structure–function correlations to be made based on continuously updated data.

**Types and numbers of mutations** There are currently 417 independently occurring dominant keratin mutations logged in the intermediate filament mutation database (http://www.interfil.org). Of these, 381 (~91.5%) are missense mutations, 28 (~6.5%) are small in-frame insertion/deletion mutations and 6 (~1.5%) are intronic splice site defects leading to larger in-frame deletions. Two out-of-frame deletions have also been reported affecting the tail domain of K1. Examples of the relative positions of the reported mutations in keratins are summarized (Fig 1). At the protein level, the functional consequences of these molecular defects are similar in that they all lead to expression of mutant polypeptides at normal or near-normal levels carrying amino acid substitutions, deletion of one or more amino acids, or in a few cases, insertion of foreign amino acids. All of these defective proteins are able to form heterodimers with the wild-type keratin partner protein and thereby integrate into the keratin network where they cause weakening or collapse of the cytoskeleton through the mechanism of dominant-negative interference. An example of transient expression of a K16 missense mutation from a PC-1 family, L124R, in cultured epithelial cells is shown (Fig 2), illustrating the complete collapse of the endogenous keratin cytoskeleton.

**Epidermolysis bullosa simplex (EBS)** as a prototypic keratin disease The inherited skin blistering disorder EBS was the first keratin disorder to be discovered in the early 1990s and is considered one of the most severe keratin disease phenotypes (Bonifas et al., 1991; Coulombe et al., 1991; Lane et al., 1992). Histologically, there is cytolysis and blistering through the basal cell layer of the epidermis in all subtypes of EBS and the major forms of the disease are caused by mutations in either keratin K5 or K14, the intermediate filament proteins specifically expressed within this cell compartment (Irvine and McLean, 1999; Smith, 2003). EBS differs from the other disorders in this group in as much as there are a number of well-defined clinical subtypes with varying degrees of severity (Fine et al., 1991, 2000).

---

**Figure 1**

Schematic diagram showing the protein domain organization of type II keratins. The rod domain consists of α-helical segments 1A, 1B, 2A and 2B, separated by non-helical linkers L1, L12 and L2. The highly conserved helix initiation and termination motifs are shaded red and green, respectively. The rod domain is flanked by variable head and tail domains, V1 and V2. Within the head and tail are short homology subdomains, H1 and H2, respectively, which are largely specific to type II keratins. (A) Positions and numbers of the reported dominant mutations in K5 and K14 causing epidermolysis bullosa simplex (EBS), revealing the mutation clusters associated with severe generalized (pink) and mild, site-restricted (yellow) EBS phenotypes. Note that the helix boundary motifs are hotspots for the more severe Dowling–Meara form of EBS. (B) Positions and numbers of the reported dominant mutations in K6a, K6b, K16, and K17 causing pachyonychia congenita (PC). Note that most of the reported mutations in the classical PC-1 and PC-2 subtypes occur in the helix boundary motifs and relatively few are found elsewhere, implying that milder PC phenotypes may not always present clinically. Note also that some mutations in the helix boundary motifs lead to milder, site-restricted clinical phenotypes (yellow), specifically focal non-epidermolytic palmoplantar keratoderma (FNEPPK) (a variant of PC-1) and steatocystoma multiplex (an allelic form of PC-2). Note that the two reports so far of late-onset PC (shaded orange) occur in sites analogous to the milder EBS mutations (shown in A).
Specifically, the Dowling–Meara form of EBS (EBS-DM) is the most severe and is life-threatening in neonates. This type of EBS is characterized by very widespread, clustered blisters and is often accompanied by fairly severe palmo-plantar keratoderma. Ultrastructurally, EBS-DM can be diagnosed by the appearance of electron-dense cytoplasmic aggregates (Anton-Lamprecht and Schnyder, 1982), which by immuno-electron microscopy, have been shown to consist of abnormally folded and condensed K5 and K14 (Ishida-Yamamoto et al., 1991). In contrast, patients with the Weber–Cockayne variant (EBS-WC) normally have blisters only on the hands and feet, although they will develop blisters elsewhere with the application of sufficient mechanical trauma. The Kőbner form of EBS (EBS-K) is intermediate in severity between EBS-DM and EBS-WC, with blisters on the hands and feet and elsewhere, but without the characteristic clustering of lesions seen in EBS-DM. There is clearly room for overlap in these milder phenotypes and many clinicians do not distinguish EBS-K and EBS-WC but rather group these together under the EBS-WC title (particularly in the UK and Europe). By electron microscopy, the basal cell keratin filaments appear to be more or less normal in the milder forms of EBS. Another mild variant of EBS is the form associated with mottled pigmentation (EBS-MP). This is again very similar to EBS-WC in severity but can be readily distinguished clinically by the pigmentary changes specific to this variant.

**Genotype–phenotype correlation in EBS and other disorders**

From the many reported EBS mutations in K5 and K14 (128 at the time of writing), about six mutation “hot-spots” have emerged, as illustrated (Fig 1). The highly conserved helix boundary motifs of both proteins are the major hotspots that were identified early on in EBS-DM patients. These short (~20 amino acids) sequence motifs that mark the beginning and end of the keratin rod domain, are known to be involved in critical molecular overlap interactions in the assembly of keratin filaments from their small heterodimer subunits. Thus, mutations in these motifs interfere with the early stages of filament elongation and result in negligible or very poor filament assembly. The unpolymerized keratins form dense cytoplasmic aggregates as seen in cultured cells (Fig 2) or in the skin of EBS-DM patients by electron microscopy (Anton-Lamprecht and Schnyder, 1982; Ishida-Yamamoto et al., 1991).

The milder EBS-K and EBS-WC variants are associated with mutations in a small number of discrete sites elsewhere along the keratin molecule. Specifically, these are the H1 domain of K5; and the second half of the 1A domain, the L12 domain, and central 2B domain of both proteins. One particular mutation, P25L in the V1 domain of K5, has been reported in all cases of EBS-MP characterized to date. The mechanism by which this particular mutation leads to abnormal pigmentation remains unknown. Because they do not interfere with the elongation process during filament assembly, all these milder mutations allow formation of normal-looking keratin filaments but result in structurally weakened filaments that break under mild mechanical stress. These weak filaments, however, provide cells with more structural resilience than the completely aggregated keratins found in EBS-DM keratinocytes and so it requires more trauma to induce cell lysis and blistering, i.e., milder, EBS phenotypes restricted to high trauma sites such as hands and feet.

Therefore, based on the site of a mutation in K5 or K14, it is largely possible to predict the resulting phenotype (Irvine and McLean, 1999; Smith, 2003). This is particularly useful in the diagnosis of infants with EBS where the clinical differences between the subtypes are less obvious. Molecular analysis in these cases can inform families and their clinicians at this early stage as to what the likely longer-term outcome will be, so that appropriate interventions can be taken. In the other keratin disorders, genotype–phenotype correlation is less of an issue and the vast majority of reported mutations occur in the helix boundary motif regions.
PC is one exception where some milder related phenotypes are now known, as will be discussed in detail below. Another exception is the case of K1, in which certain mutations have been shown to produce milder site-restricted phenotypes, compared with the severe generalized epidermolytic hyperkeratosis phenotype normally associated with K1 and K10 mutations (bullous congenital ichthyosiform erythroderma). Some of these are missense mutations in the helix boundary motifs—such as I479T in the helix termination motif of K1, which has been associated with mild ichthyosis-like phenotype in some cases (Sybert et al., 1999) and epidermolytic palmoplantar keratoderma alone in other families (Terron-Kwiatkowski et al., 2004). It is not clear why this mutation, which occurs in a region normally associated with severe phenotypes in other keratins, results in a mild phenotype where K1 is concerned. Larger in-frame deletions in K1 have also been shown to result in epidermolytic palmoplantar keratoderma—known as “K1 keratoderma” (Hatsell et al., 2001; Terron-Kwiatkowski et al., 2002). It may be that these larger deletions, particularly those removing substantial portions of the boundary motif sequences, render the mutant polypeptide less able to assemble and are therefore less disruptive. This seems to be the case with a similar deletion in K16 leading to a mild PC-related phenotype, discussed below (Smith et al., 2000).

**Emerging genotype–phenotype trends in PC**

As is the case with most non-EBs keratin mutations, the vast majority of the pathogenic defects in the PC-1 keratins (K6a, K16) and the PC-2 keratins (K6b, K17) occur in the helix boundary motifs of these proteins, including the novel and recurrent mutations reported elsewhere in this issue (Smith et al.). The phenotypes in these cases are fairly consistent—typical severe nail dystrophy and other ectodermal features. In keeping with the other keratins, the majority of these mutations are missense changes, with a smaller number of in-frame insertions or deletions.

Two milder phenotypic variants of PC-1 and PC-2 have been recognized, FNEPPK and steatocystoma multiplex, respectively, as discussed above. Certain K16 defects have been reported to cause FNEPPK: these are the missense mutations N125S and R127C in the helix initiation motif (Shamsher et al., 1995); and a complex 24 bp deletion in the helix termination motif of the K16 polypeptide—a deletion of 23 bp and a separate 1 bp deletion nearby (Smith et al., 2000). Elsewhere in this issue, two independent occurrences of the N125S mutation are reported (Smith et al.). In one of these families, this mutation is again associated with a milder phenotype consisting of FNEPPK with trivial nail changes. The other family, however, presented with a more classical PC-1 phenotype including hypertrophic dystrophy of most nails. Therefore, N125S appears to be more often, but not always, associated with the milder phenotype. These cases emphasize the importance of not basing hard and fast conclusions about genotype–phenotype correlations on small numbers of families. Interestingly, a second family with the FNEPPK phenotype studied by Smith et al. (this issue), carried the R127C mutation in K16, which has also been reported previously in association with this milder presentation (Shamsher et al., 1995). The equivalent mutation in K14 (R125C) consistently results in the severe EBS-DM phenotype (Coulombe et al., 1991; http://www.interfil.org). Thus, there may be more genetic or environmental modifiers at play in PC compared with EBS.

The second type of mutation associated with FNEPPK, the complex 24 bp deletion involving the helix termination motif of K16 (Smith et al., 2000), is similar in its effect to the deletions seen in K1 that produce milder, body site restricted phenotypes (Hatsell et al., 2001; Terron-Kwiatkowski et al., 2002). In the case of this K16 mutation, protein expression studies in cultured cells showed that the deletion appears to have less detrimental effects on the endogenous cytoskeleton compared to missense mutations (Smith et al., 2000). This is consistent with the idea that these types of mutations render the mutant protein less able to interact with and therefore disrupt wild-type keratin subunits.

In the case of steatocystoma multiplex, a milder, phenotypically restricted “relative” of PC-2, only a few mutations have been reported to date, all in the helix initiation motif of K17, specifically, N92H, R94H, and R94C (Smith et al., 1997; Covello et al., 1998; Wang et al., 2001). What is particularly interesting here is the fact that mutation R94C was reported in one family causing the classic PC-2 phenotype but in two other unrelated kindreds, gave the steatocystoma multiplex phenotype without significant nail changes (Covello et al., 1998; Wang et al., 2001). In each case, the phenotype was consistent within the family. This clearly demonstrates that there are genetic background effects *i.e.* modifying genes or other factors, that can alter the severity of the PC phenotype. Similar phenotypic variation has been observed in most or all of the known keratin disorders but the identity of the modifying factors underlying this epiphenomenon remains unknown. This also has important consequences for making genotype–phenotype correlations and predictions about future disease patterns or severity based on molecular data. For example, the equivalent mutation in K14 (R125C) to the steatocystoma defect R94C consistently produces the EBS-DM phenotype (Coulombe et al., 1991). Therefore caution should be exercised when making prognostic predictions based on a relatively small sample of mutation-phenotype data.

To date, no K6a mutations have been reported in FNEPPK families, although we have looked for these without success (Smith and McLean, unpublished observations). Similarly, no K6b mutations have been seen in steatocystoma multiplex. Since type II keratin mutations accurately phenocopy the respective type I mutations in PC-1 and PC-2, there is no obvious reason why K6a and K6b mutations would not occur in association with milder presentations. Again, the overall numbers of mutations in these genes are somewhat small at this stage.

**Mild mutations may modulate age of onset in PC**

Normally, PC is evident in the first months of life, either as skin blistering, hyperkeratosis or nail dystrophy. One clinical variant of PC has been described where there is a relatively late age of onset, described as PC-tarda. Recently, the molecular defects underlying two cases of PC-tarda have been identified. In the first case, a missense mutation, K354N, located in the central 2B domain of K16, was identified in a case of late-onset PC-1 (Connors et al., 2001). In the second case, reported very recently, late-onset PC-2
was found to be caused by a missense mutation in the second half of the 1A domain of K17, N109D (Smith, 2004; Xiao et al, 2004). Significantly, both of these mutations occur in regions equivalent to the hotspots of K5 and K14 that are associated with the mild EBS-WC/EBS-K phenotypes (Fig 1). Thus, two mutations of a type that would result in a mild EBS phenotype if they occurred in K5 or K14, here led to delayed onset PC. As stated above, it is not advisable to draw firm genotype–phenotype conclusions based on very small numbers of mutations, however, these results perhaps point to an emerging trend where mutations in L12 domain of K6a, K6b, K16, and K17 or the H1 subdomain of the type II keratins, might be predicted to give late-onset phenotypes. Further mutation analysis should confirm these trends in the future.

Keratin polymorphisms and modifier genes Recently, there has been a report of a second modifying mutation within a keratin gene (Yasukawa et al, 2002). In this study, a family presented with several members affected by autosomal dominant EBS, some of whom were mildly affected, whilst others quite severely affected by skin blistering. All the affected persons, regardless of severity, carried a mutation in the 1A domain of K5, E170K. The more severely affected individuals also carried a second mutation on the other K5 allele, E418K in the center of the 2B domain. Interestingly, a number of people in the family carried only the E418K mutation and had no skin blistering phenotype whatsoever, even though this mutation was shown to produce a low level of keratin aggregation when expressed in cultured cells (Yasukawa et al, 2002). Neither of these mutations were detected in 100 ethnically matched chromosomes and so were not common sequence variants in the population. Thus, the E418K mutation acts as a polymorphism in the heterozygous state but is able to exacerbate a mild EBS phenotype when in the compound heterozygous state. It is not known what the effect of a homozygous E418K mutation would be in humans but it is quite possible that this might produce recessive EBS, similar to other reports in the literature (Hovnanian et al, 1993). This study points the way towards understanding at least some of the factors involved in generating phenotypic variation within a kindred with a given mutation. Very often, genetic testing laboratories do not fully screen genes but halt the screening process as soon as a mutation is detected and confirmed. This is a particularly common practice in keratin gene screening since analysis generally starts with the most prevalent mutation hotspots (Fig 1) and only when these are negative, are the other exons or other keratin genes screened. Thus, important modifying factors within a keratin pair can easily go unnoticed.

To date, there has been no formal publication of non-pathogenic polymorphisms in the keratin genes causing PC, although from our own screening program, we have identified a number of mainly silent polymorphisms in these genes, e.g., A161A and E165E in the K6a gene (Smith, unpublished data). Furthermore, there are many single nucleotide polymorphisms in the genome database involving these genes, several of which affect exons (http://genome.ucsc.edu). Therefore, there is some potential for keratin polymorphisms acting as phenotypic modifiers in PC and full sequencing of K6a, K6b, K16, and K17 genes might help identify these genetic modifiers within families presenting with highly variable phenotypes (see Leachman et al, this issue).

Conclusions

There are four main conclusions to draw from the study of mutations in PC. Firstly, and consistently to date, mutations in K6a or K16 give rise to the PC-1 phenotype whereas K6b or K17 mutations result in the more complex PC-2 phenotype. Secondly, certain mutations in K16 can give rise to the milder FNEPPK phenotype and similarly, a subset of mutations in K17 cause steatocystoma multiplex. Thirdly, some keratin mutations can produce mild or severe phenotypes in different families/ethnic backgrounds, pointing to the existence of unknown genetic and/or environmental modifying factors and a need for caution in predicting phenotypes purely from mutational data. Finally, as well as milder site-restricted phenotypes, some mutations, particularly those similar to the milder EBS mutations, may lead to delayed onset forms of PC.

There is still much to learn about the genetics of PC. In our experience, about 10% of PC patients referred to our laboratory for diagnostic testing lack detectable mutations in any of the four appropriate keratins. Although we were able to account for some of these missing mutations by the misdiagnosis of Clouston syndrome patients who carry connexin 30 defects but resemble PC clinically (van Steensel et al, 2003), there still remain a group of PC patients who lack mutations. Some of these may have larger genomic deletions, duplications or other rearrangements that go undetected using the normal exon-by-exon PCR mutation detection strategies. Since the keratins involved in PC have multiple pseudogenes and/or isoforms, these types of mutations may not be uncommon, as is the case for other clusters of highly homologous genes, such as the globin genes (Cooper and Krawczak, 1993). Southern blot analysis or the recently developed MLPA technique (multiplex ligation-dependent probe amplification) could be used to detect these types of defects (Schouten et al, 2002). Activation of a partly translatable keratin pseudogene carrying missense mutations or other sequence changes is another possibility, if a genomic rearrangement placed such a gene downstream of an active keratin promoter. There is also the possibility that mutations in other K6 isoforms, such as K6hf or others (Takahashi et al, 1995; Winter et al, 1998), may cause variants of PC that phenotypically resemble PC-1 or PC-2. The recent completion of the human genomic sequence across the type I and II keratin gene clusters should facilitate identification of these more cryptic types of PC mutation and may also aid in elucidation of some of the genetic modifying factors involved. Overall, this should lead to a fuller understanding of the molecular basis of PC, improve diagnostic testing and ultimately, help in the rational design of therapies for this group of incurable diseases.

Materials and Methods

Software Adobe GoLive CS web design software (Adobe Systems Inc., San Jose, CA) was used to create the HTML code which
We thank the patients and clinicians who facilitated our original studies on the genetic basis of keratin diseases. Frances Smith and Irwin McLean are currently funded by a Wellcome Trust Senior Research Fellowship (W. H. I. M). Research in the Epithelial Genetics Group is also supported by The Dystrophic Epidermolysis Bullosa Research Association UK (http://www.debra.org.uk) and The Pachyonychia Congenita Project (http://www.pc-project.org).

DOI: 10.1111/j.1087-0024.2005.10205.x

Manuscript received June 9, 2005; revised June 27, 2005; accepted for publication June 28, 2005

Address correspondence to: Professor Irwin McLean, Human Genetics Unit, Ninewells Medical School, Dundee DD1 9SY, UK. Email: w.h.i.mclean@dundee.cc.uk

References


Dunnen JT, Antonarakis SE: Mutation nomenclature extensions and suggestions to describe complex mutations: A discussion. Hum Mutat 15:7–12, 2000


