INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

MAPPING OF CLOUSTON HIDROTIC ECTODERMAL DYSPLASIA

by
Zoha D. Kibar

Department of Biology
McGill University

February, 1999

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements of the degree of Doctor of Philosophy

ACNOWLEDGEMENTS AND DEDICATION

I would like to thank my supervisor, Dr. Guy A. Rouleau, for his guidance and support throughout the project. I would also like to thank the members of my supervisory committee, Drs. Vazken M. Der Kaloustian, F. Clarke Fraser and Paul Lasko, for helpful advice. I would like to acknowledge all the participating clinicians and families who made this study possible.

I would also like to thank my friends and colleagues for their support: Patrick Dion, Marie-Pierre Dubé, Claudia Gaspar, Heidi Howard, Mehrdad Jannatipour, Ronald Lafrenière, Claude Marineau, Jillian Parboosing, Albéna Pramatarova, Daniel Rochefort, Vijayalakshmi Shanmugam, André Toulouse and Gustavo Turecki. Special thanks to Ron, Mehrdad, Patrick, Viji and André for making work more interesting and enjoyable.

This study was supported by grants from the Canadian Dermatology Foundation and the National Foundation for Ectodermal Dysplasias. Dr. Guy.A.Rouleau is supported by the Medical Research Council of Canada.

This work is dedicated to my parents, Nouhad and Doureid Kibar, for their love and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Clouston hidrotic ectodermal dysplasia (HED) is an autosomal dominant skin disorder that is characterized by nail dystrophy, hair defects and palmoplantar hyperkeratosis. This condition has been described in families of various ethnic origins but is particularly common in the French Canadian population. Using linkage analysis in eight French Canadian families segregating HED, we mapped the HED gene to the pericentromeric region of chromosome 13q with a combined two-point lod score of 8.12 at zero recombination from the marker D13S175. Haplotype analysis allowed us to define D13S143 as the telomeric flanking marker for the HED candidate region. We tested five genes that map to this region, connexin 26, connexin 46, fibroblast growth factor 9, zinc-finger *ZNF198* and α tubulin *TUBA2*, for involvement in HED by PCR-SSCP analysis. No mutation specific to HED was found in any of them suggesting that they most likely are not defective in this disease.

To facilitate the identification of the HED gene, we constructed a radiation hybrid (RH) map of 48 loci surrounding the HED locus on chromosome 13q. This map integrates 3 genes (*TUBA2*, *GJβ2* and *FGF-9*) and 18 ESTs with 27 markers including 19 polymorphic loci. A major inconsistency in order involving a reversed interval of six loci was found between our RH map and a YAC contig established in the region. We used Fiber-FISH and FISH on interphase nuclei to confirm our order. To refine the localization of the HED gene, we isolated eight new chromosome 13q polymorphic $(CA)_n$ markers and used seven of them along with three others in genetic analysis of a multiethnic group of 29 HED families. We demonstrated genetic homogeneity in HED in four families of French, Spanish, African and Malaysian origins and showed evidence for a strong founder effect in families of French Canadian origin. Recombination mapping placed the HED gene in a 2.4 cM region flanked by D13S1828 proximally and D13S1830 distally. Multipoint linkage and linkage disequilibrium analyses finely mapped the HED gene at 0-0.08 cM telomeric to D13S1835. These studies will greatly facilitate the physical mapping and positional cloning of the HED gene.

RÉSUMÉ

La dysplasie ectodermique hidrotique (HED) de type Clouston est une maladie cutanée à transmission autosomale dominante. Elle est caractérisée par une dystrophie des ongles, une hypotrichose et une hyperkératose palmo-plantaire. Cette condition a été décrite chez des familles de diverses origines ethniques mais est particulièrement fréquente dans la population Canadienne-Française. Nous avons pu cartographier le gène HED dans la région péricentromérique du chromosome 13q avec un two-point lod score combiné de 8.12 sans recombinaison à partir du marqueur D13S175 au moyen d'une analyse de liaison chez huit familles canadiennes-françaises atteintes. Des analyses d'haplotype ont ensuite permis d'identifier le marqueur D13S143 comme borne télomérique de la région candidate pour HED. Par analyse de PCR-SSCP, nous avons évalué l'implication potentielle de cinq gènes cartographiés dans cette région, soient connexin 26, connexin 46, fibroblast growth factor 9, zinc-finger *ZNF198* et α tubulin *TUBA2*. Aucune mutation spécifique à HED n'a été détectée, suggérant qu'aucun de ces gènes n'est impliqué dans la maladie.

Afin de faciliter l'identification du gène HED, nous avons établi une carte d'hybrides de radiation (RH) de 48 loci environnant le locus HED sur le chromosome 13q. Cette carte intègre 3 gènes (*TUBA2*, *GJβ2* et *FGF-9*), 18 ESTs et 27 marqueurs incluant 19 loci polymorphes. En comparant notre carte RH avec un contig de YAC établi dans la région, nous avons relevé une discordance majeure impliquant un intervalle inversé de six loci. Nous avons utilisé le Fiber-FISH et le FISH sur des noyaux en interphase afin de confirmer notre ordre. Afin de raffiner la localisation du gène HED, nous avons isolé huit nouveaux marqueurs polymorphes $(CA)_n$ sur le chromsome 13q. Nous avons utilisé sept d'entre eux en plus de trois autres lors de l'analyse génétique d'un groupe multiethnique de 29 familles HED. Nous avons démontré l'homgénéité génétique dans HED chez 4 familles d'origines Française, Espagnole, Africaine et Malaysienne et démontré la présence d'un effet fondateur important chez les familles d'origine Canadienne-Française. La cartographie de recombinaison localise le gène HED dans une région de 2.4 cM flanquée par les marqueurs D13S1828 (proximal) et D13S1830 (distal). Les analyses de liaison multipoint et de déséquilibre de liaison localisent le gène HED

entre 0 et 0.8 cM du côté télomérique de D13S1835. Ces études vont grandement faciliter la cartographie physique et le clonage positionnel du gène HED.

CHAPTER 1

Introduction

1.1. Introduction to ectodermal dysplasias

1.1.1. Definition and clinical features

The term "ectodermal dysplasia" was first used by Weech in 1929 to describe a small group of diseases in which the affected tissues are of ectodermal origin, the defects are developmental and hereditary transmission is well established. At that time, the term was proposed to describe mainly the Christ-Siemens-Touraine's (CST) syndrome or X-linked anhidrotic ectodermal dysplasia (EDA), the most common of the ectodermal dysplasias characterized by sparse hair (hypotrichosis), abnormal or absent teeth (hypodontia or anodontia) and inability to sweat due to lack of sweat glands (anhidrosis) (Solomon and Keuer, 1980). Later, Clouston described a new form of hereditary EDs which is hidrotic, in contrast to CST which is anhidrotic. This hidrotic ED (HED) segregates as an autosomal dominant condition with nail dystrophy, defects of the hair and palmoplantar hyperkeratosis. In his papers, Clouston presented four new cases of the anhidrotic type (CST) and described in details the hidrotic type in a large French Canadian family (Clouston, 1929; 1939).

Many different clinical symptoms are encountered in EDs. The hair can be sparse or completely absent, abnormally distributed, thin and fragile with slow growth rate, abnormally shaped (e.g. presence of longitudinal fissures) or hypopigmented. The teeth can have an abnormal shape (e.g. conical, large and bulbous crowns), be present at birth (natal teeth), be supernumerary, be smaller (microdontia) or be congenitally absent (hypodontia). The nails can be malformed with slow growth rate, hypopigmented, discolored or completely absent. The sweat glands can be regionally or totally absent, hypoplastic, misplaced or dysfunctional. EDs can also be associated with general

alterations of the skin (e.g. blistering, hyperpigmentation), disorders of nasal and oral mucosa, mental retardation, endocrine disturbances, hearing loss, ocular abnormalities, limb malformations as well as other developmental defects (Freire-Maia, 1971; Holbrook, 1988).

To date, EDs constitute a clinically and genetically heterogeneous group of more than 150 inherited disorders affecting tissues of ectodermal origin. Classification and delineation of this complex nosologic group were proposed by Freire-Maia who arbitrarily categorized them into two groups. Group A includes all the conditions with defects in a least two of the following structures-hair, teeth, nails and sweat glands, with or without other defects. Group B includes conditions with defects in only one of these four structures plus at least another ectodermal defect (Freire-Maia, 1971; Pinheiro and Freire-Maia, 1994). Clouston HED was classified by Freire-Maia and Pinheiro (1984) as Group A subgroup 1-2-3 (hair, teeth, nail) because of the contradictory findings of associated dental anomalies. However, later studies demonstrated the absence of significant morphologic dental abnormalities in HED and this disease was suggested to be classified as a hair-nail (1-3) dysplasia (Hassed *et al.*, 1996).

1.1.2. Molecular defects in EDs

The hereditary nature has been established for more than 100 EDs with either autosomal dominant or autosomal recessive or X-linked mode of transmission (Pinheiro and Freire-Maia, 1994). However, the underlying genetic defects have been identified only for few of them.

The largest group of EDs for which the genes have been identified are caused by mutations in keratin genes (Mclean and Lane, 1995). Keratins belong to the group of intermediate filaments that along with actin filaments and microtubules form an extensive cytoskeleton in all epithelial cells (Fuchs, 1995). Examples of mutated soft (epidermal) keratins in EDs include: the basal K5 and K14 in epidermolysis bullosa simplex (EBS) characterized by skin fragility and basal cytolysis (Coulombe *et al.*, 1991; Lane *et al.*, 1992), the suprabasal K1 and K10 in epidermolytic hyperkeratosis (EH) similar to EBS but with cytolysis occurring in the suprabasal layer of the epidermis (Rothnagel *et al.*, 1992), the palm and sole specific K9 in diffuse epidermolytic palmoplantar keratoderma typified by palmoplantar skin blistering (Reis *et al.*, 1994) and the hyperproliferative K16 and K17 in pachyonychia congenita characterized mainly by hypertrophic nail dystrophy (McLean *et al.*, 1995). To date, only one example of mutated hard trichocyte (hair) keratins has been described in the literature: hHb6 and hHb1 in monilethrix that is a rare dominant hair disease characterized by dystrophic alopecia of varying size and rough follicular papules (Winter *et al.*, 1997a; 1997b). Mutations in a keratin- associated protein, loricrin, were detected in autosomal dominant erythrokeratoderma characterized by widespread erythematous plaques associated with palmoplantar keratoderma (Ishida-Yamamoto *et al.*, 1997).

Another group of EDs are caused by mutations in structural components of the dermoepidermal basement membrane zone and the adherens junctions of epithelial cells. Mutations in the 180 kD bullous pemphigoid antigen/type XVII collagen, plectin and integrin $\alpha 6\beta 4$ underlie three hemidesmosomal forms of EB, generalized atrophic benign EB (McGrath *et al.*, 1995), EB associated with late-onset muscular dystrophy (Pulkkinen

et al., 1996) and EB associated with pyloric atresia (Vidal *et al.*, 1995) respectively. Mutations in collagen VII cause dystrophic EB with scarring forms (Christiano *et al.*, 1993) and mutations in laminin γ2 chain result in junctional EB (Pulkkinen *et al.*, 1994). Mutations in plakophilin 1 underlie the ectodermal dysplasia/skin fragility syndrome characterized by skin blistering and desquamation associated with short and sparse hair and thick dystrophic nails (McGrath *et al.*, 1997).

In addition to abnormalities of structural proteins, defects in enzymes involved in the normal physiology of the skin have been described in EDs. The enzyme steroid sulfatase involved in lipid metabolism in the stratum corneum is defective in recessive X-linked ichthyosis characterized by extremely scaly skin and corneal opacities (Shapiro *et al.*, 1989). Mutations in transglutaminase that catalyzes cross-linking of proteins of the CE cause lamellar ichtyosis characterized by excessive thickening of the stratum corneum giving the skin a scaly appearance (Russel *et al.*, 1995).

Two types of EDs have been associated with defects in genes presumably involved in the developmental process of the ectodermal appendages. Alopecia universalis (AU), a form of hereditary hair loss, is caused by mutations in the human hairless gene that encodes a putative single zinc finger transcription protein believed to regulate catagen remodeling in the hair cycle (Ahmad *et al.*, 1998). X-linked anhidrotic ectodermal dysplasia (EDA) results from mutations in a novel protein with a transmembrane domain and collagen-like helix forming structures. The authors suggest that the EDA gene might belong to a novel class with a role in epithelial-mesenchymal signaling, as either a cellular receptor or coreceptor or as a membrane-bound ligand, or

that it might function in cell adhesion or migration through interaction with the extracellular matrix (Kere *et al.*, 1996; Monreal *et al.*, 1998).

1.2. Clouston hidrotic ectodermal dysplasia

1.2.1. History and mode of inheritance

Clouston hidrotic ectodermal dysplasia (HED) is a rare autosomal dominant disorder transmitted with complete penetrance and variable clinical expressivity (Williams and Fraser, 1967). The first cases of HED were described by Nicolle and Hallipré in a French Canadian family in 1895 (George and Escobar, 1984). However, the hereditary nature of HED was established only in 1929 by Clouston who described this disorder in five consecutive generations of a large French Canadian family living in Québec (Clouston, 1929). He concluded that the condition is transmitted as an autosomal dominant trait. In a later investigation of the same family, Williams and Fraser (1967) confirmed this conclusion and found no deviations from an autosomal dominant pattern of inheritance. To date, most of the families segregating this condition are of either French or French-Canadian origin, suggesting a common ancestor. Occurrence of HED in families of other ethnic origins such as African American (McNaughton and Pierson, 1976), Chinese (Rajagopolan and Hai Tay, 1977), Japanese (Ando *et al.*, 1988), British (Patel *et al.*, 1991), Indian (Radhakrishma *et al.*, 1997) and Irish-Scottish (Taylor *et al.*, 1998) has also been reported.

1.2.2. Clinical features

The main features of HED are nail dystrophy, defects of the hair and hyperkeratosis of the palms and soles. The nail changes can occur in varying degrees of severity, ranging from thickening and discoloration to the presence of marked longitudinal striations, deep pigmentation and detachment from the nail bed. They are more susceptible to paronychial infections. The hair is typically thin and sparse on the eyebrows and body but the anomalies can range from brittleness and slow growth rate to total alopecia. The skin generally appears hyperkeratotic and dry despite the presence of normal sweat glands. The skin of the palms and soles is the most severely affected, with a rough and easily cracked appearance (Der Kaloustian and Kurban, 1979; Escobar *et al.*, 1983; Ando *et al.*, 1988). Variable and infrequent features reported with HED include thickening of the skull bones, tufting of the terminal pharynges, brownish pigmentation of the skin over the knuckles, elbows, knees, axillae, areolae and pubic areas (Williams and Fraser, 1967; Escobar *et al.*, 1983), oral lesions (George and Escobar, 1984), ocular abnormalities (Hazen *et al.*, 1980), sensorineural deafness, epilepsy, mental retardation and skeletal anomalies such as polydactyly and syndactyly (Pierard *et al.*, 1979; Worobec-Victor *et al.*, 1988). In contrast to the X-linked anhidrotic type of ectodermal dysplasia (EDA), facial appearance, teeth growth and sweating are normal in HED (Der Kaloustian and Kurban, 1979).

1.2.3. Biochemical and biophysical studies

Most pathological studies conducted in HED are of the defective hair. Normally, the hair fiber consists of a thin cuticle enclosing a central cortex made up of alpha-helical hard keratinous fibers embedded in an interfilamentous globular matrix. The matrix is

highly rich in cysteine, relatively rich in proline and serine and relatively poor in phenylalanine and tyrosine and interacts with the keratin fibers mainly though disulfide bonds. In HED, biophysical and biochemical studies suggest a depletion of matrix protein related to disruption of disulfide bond formation in the remaining keratin. The hair is thin with abnormal swelling capacity, reduced tensile strength, reduced birefringence under polarized light, and looser and disorganized structure by light microscopy. Despite this distorted hair architecture, X-ray diffraction analysis showed an intact α-helical structure of the keratin fibers. Chemical alterations of the hair include decreased levels of disulfide bonds, cysteine, proline and serine and increased amounts of reactive SH groups, tyrosine and phenylalanine. The epidermal soft keratins had a normal amino acid composition (Gold and Scriver, 1972). The diminished cysteine content in HED was also shown in a biochemical study of keratin of the defective nails in HED (Giraud *et al.*, 1977). Ultrastructural findings in HED revealed the presence of longitudinal grooving of the affected hair with an abnormal cross-sectional shape and disorganized fibrous cortex. The hair is smaller in diameter with increased desquamation of the cuticle that could be due to a reduction of disulfide bonds (Escobar *et al.*, 1983).

Light microscopic examination of biopsies of the affected palms and soles of HED patients showed marked hyperkeratosis of the epidermis with normal appearance of the epidermal cells and no epidermolytic changes. Electron microscopy results in the same study revealed an increase in the number of desmosomes in the thickened stratum corneum of the epidermis suggesting that hyperkeratosis could be due to the delayed desquamation of this layer (Ando *et al.*, 1988).

1.2.4. Possible underlying defect in HED

1.2.4.a. Molecular abnormality of keratin

The abnormal findings in studies of the defective hair and nails in HED suggest a defect in keratin as a possible causative pathogenic mutation. Keratins are the major structural proteins of the epidermis and its appendages. Based on amino acid sequence homology and electrophoretic properties, they can be divided into two groups: type I acidic and type II basic. Type I keratins include twelve epithelial soft proteins, K9-K20, and seven hair/nail keratins. Type II keratins include eight epithelial keratins, K1-K8, and four hair/nail keratins (Rogers *et al.*, 1997). All keratins share a common secondary structure: a central conserved α-helical rod domain flanked by non-helical and more variable amino- and carboxy-terminal domains (called head and tail domains respectively). While the rod domains define the molecular alignments in the keratin filaments, the head and tail domains have been implicated in interkeratin filament association and interaction of keratin filaments with other proteins in the cell conferring specialized functions to the particular cell. (Steinert and Roop, 1988; Steinert, 1993; Yu *et al.*, 1993; Steinert *et al.*, 1994).

The keratin intermediate filaments (KIFs) are embedded in a matrix composed of a group of proteins known as intermediate filament associated proteins (IFAPs). In the hair, IFAPs consist of several groups: the high-sulfur (HS) group, the ultrahigh-sulfur (UHS) group, the high-glycine/tyrosine group (HGT) and trichohyalin. The HS, UHS, HGT groups interact with KIFs mainly though disulfide bonds. Trichohyalin, only found in the medulla and the inner root sheath of the hair, is crosslinked to KIFs through

isopeptide bonds by the enzyme transglutaminase (Powell and Rogers, 1990a; Rogers *et al.*, 1991).

Epidermal soft KIFs genes are found in clusters mapped to two different chromosomes: chromosome 17q12-21 (type I) and chromosome 12q11-12 (type II) (Lessen *et al.*, 1988; Rosenberg *et al.*, 1988). Human hair keratin genes are co-localized with the epidermal keratin genes on these chromosomes (Rogers *et al.*, 1995; Bowden *et al.*, 1998). Genes encoding the UHS proteins have been localized to two different clusters on chromosome 11 in humans (Mackinnon *et al.*, 1991). The trichohyalin gene has been mapped to a cluster of genes coding for other epidermal structural proteins localized at 1q21 (Lee *et al.*, 1993).

The keratin genes are expressed in a specific spatial and temporal pattern in the different epithelial tissues at different stages of development and differentiation suggesting the presence of a highly complex genetic regulation (Steinert and Roop, 1988; Powell and Rogers, 1993; 1997). One type I keratin combines with a specific type II partner in a tissue- and differentiation-specific manner. For example, K5 and K14 are co-expressed in all basal cells, K1 and K10 in the stratified epithelial epidermal cells and so on. This co-expression of gene pairs suggests a coordinate control (O'Guin *et al.*, 1990). Keratin gene expression is mainly regulated at the transcriptional level. Some different regulatory elements have been recognized in their promoter regions and introns that may be responsible for their differential expression (Blessing *et al.*, 1987; 1989; Jiang *et al.*, 1991; Powell *et al.*, 1991; 1992; Powell and Beltrame, 1994). In addition, translational and posttranslational regulation of keratin genes, namely phosphorylation, have been implied in some studies (Tyner and Fuchs, 1986; Kulesh *et al.*, 1989; Skalli *et al.*, 1992).

Certain environmental factors have also been shown to influence cytokeratin synthesis; they include hormones, growth factors and vitamin A related compounds (Blumenberg *et al.*, 1992).

Mutations in keratin genes have been described in many ectodermal dysplasias stressing their importance in maintaining structural integrity and providing mechanical strength to the epithelial cells (reviewed in Steinert and Bale, 1993; Fuchs, 1996a; 1996b; 1997). Most of these mutations are clustered in the highly conserved amino- or carboxy-end of the rod domain that are critical for the proper packing of keratin subunits into 10 nm keratin filaments. They frequently occur in the same residue position on the different chains. Recent transgenic mouse studies have further demonstrated the importance of proper assembly of the keratin filaments in epithelial cells. Overexpression of a sheep wool keratin gene in transgenic mice disrupts the normal keratin protein composition and leads to weakening and premature loss of hair fibrils (Powell and Rogers; 1990b). Transgenic mice expressing a mutant K14 develop EBS (Vassar *et al.*, 1991) and those expressing a mutant K10 develop EH (Fuchs *et al.*, 1992). Overexpression of K16 in transgenic mice causes abnormalities in the terminal differentiation in the outer root sheath of the hair and in the epidermis (Takahashi *et al.*, 1994).

In summary, there is ample evidence that mutations in keratin genes are excellent candidates for HED.

1.2.4.b. Developmental error

Pierard et al (1979) suggested that ectodermal dysplasias might be the result of a disturbed control mechanism in the mesoderm-ectoderm interaction during

morphogenesis of the ectodermal tissues. Having almost no data on the fetal tissues at the different stages of development in ectodermal dysplasias, Holbook (1988) discussed this hypothesis based on the abnormal findings of the epidermis and its appendages in affected adults. He reviewed the normal development and differentiation of the epidermis and epidermal appendages and discussed the various factors that might interfere with these processes leading to the development of EDs.

Morphogenesis of skin appendages such as hair and nail is a very complex process whereby interactions between the epithelium and the mesenchyme result in the conversion of a flat piece of ectoderm into an epidermal appendage with unique structure and composition. The stages of this process have been well described morphologically and many studies have been conducted to investigate the underlying molecular events. Though most of these studies involve the hair follicle, all epidermal appendages are believed to follow a similar pattern of development (Holbrook, 1988). The process begins with an inductive event that involves a series of messages between the dermis and the epidermis. The initial message is sent by the dermis and instructs the epidermis to form an appendage. This is followed by a specific signal from the epidermis that instructs the dermis to form dermal papillae. Finally, a message stimulates the epidermis to form the specific appendage. Postinductive stages include initiation, elongation and differentiation of the appendage (Holbrook and Minamin, 1991; Messenger, 1993). The dermal-epidermal communication involves cell-cell contact, cell-matrix interactions and association with nerves. These different kinds of interactions involve molecules, called morphogens, that play an important role in the regulation of the development (Holbrook *et al.*, 1993; Stenn *et al.*, 1996). These morphogens include cell adhesion molecules (e.g.

neural-cell adhesion molecules, cadherins, tenascin and integrin) (Chuong *et al.*, 1991) extracellular matrix molecules (e.g. proteoglycans, glycoproteins, type IV collagen, laminin) (Couchman *et al.*, 1991; Messenger *et al.*, 1991; Couchman, 1993), growth factors (epidermal, transforming and fibroblast growth factors) (Moore *et al.*, 1991; Peus and Pittelkow, 1996), gap junction molecules or connexins (Goliger and Paul, 1994), hormones and enzymes. By investigating the presence of these molecules in the microenvironment of the developing skin appendage and studying their expression pattern to see if they are tissue- or stage-specific, many of these molecules have been shown to be involved in the morphogenesis of the follicle. One could speculate that any defect involving these molecules at the dermal-epidermal junction could lead to abnormal development of the epidermal appendages.

1.3. Rationale and goal of thesis project

The gene responsible for HED has not been identified yet. The goal of this project was to map the HED gene by linkage analysis and to refine its localization to a region amenable to positional cloning strategies. This includes the following steps:

1) mapping the HED gene by linkage analysis

The most critical element for genetic analysis of HED is the availability of DNA from large families segregating this disease. When this work was started in January 1994, five French Canadian families segregating HED were collected. They included 19 affected among a total of 41 individuals. Additional families needed to be collected to increase the number of informative meioses for linkage analysis. By the end of September 1994, we had eight HED families that included 44 affected among a total of 80

individuals. By the end of this project, we had 29 HED families consisting of 287 subjects and including 172 affected. Most of the clinical data and the blood samples were collected in collaboration with Dr. Vazken Der Kaloustian at the division of Medical Genetics at the Montreal's Children Hospital.

Since the basic defect underlying HED was suggested to be a molecular abnormality of keratins, the HED families were to be typed first with informative polymorphic markers spanning chromosomal regions known to harbor the genes coding for keratins or their associated proteins on chromosomes 1, 11, 12 and 17. Next, they were to be typed with markers spanning chromosomes 14, 15 and 21 that have been suggested to be the sites of more distantly related keratin genes. If no linkage was found between HED and any of these candidate loci, a whole genome scan was to be initiated using well-mapped Genethon polymorphic markers. The genotyping data was to be analyzed using the MLINK and LINKMAP programs of the LINKAGE 5.1 package (Lathrop *et al.*, 1984; 1985).

2) testing candidate genes in the region for involvement in HED

Once the HED gene is localized to a chromosomal region, genes mapping to this region were to be tested for involvement in HED. They were to be screened for the presence of pathogenic mutations in HED patients by PCR-based single-strand conformation polymorphism (SSCP) (Orita *et al.*, 1989). This method detects migrational differences of PCR products on a non-denaturing polyacrylamide gel caused by sequence differences in the mutant DNA strand. Any band alteration is analyzed in 100 normal

controls and sequenced if necessary to determine its exact nature (pathogenic or normal polymorphism).

In addition, expressed sequence tags (ESTs), publicly available in the DNA sequence databases and mapping to the candidate region, were to be screened to identify homologous sequences using the Blast family of tools (Blastn, Blastp, Blastx and tBlastn) (Altschul *et al.*, 1990). Any EST giving homology to genes involved in the development and normal physiology of the skin or its appendages would be further analyzed and investigated for involvement in HED. This requires the isolation of the corresponding cDNA from commercially available cDNA libraries, characterization of the genomic structure of the gene and screening it for the presence of pathogenic mutations in HED patients as described above.

3) construction of a physical and genetic map for the HED candidate region

If none of the candidate genes were found to be responsible for HED, a positional cloning strategy is required to isolate the defective gene. As a prerequisite to this strategy, we were to construct a physical and genetic map for the candidate region. Depending on the size of the candidate region, we will use either radiation hybrid (RH) mapping or YACs or both to create a physical map of the region. RH mapping is a somatic cell genetic mapping method that orders markers based on the frequency of X-ray induced breakage between them. It provides a powerful tool that can integrate physical and genetic mapping data at a level of resolution intermediate between genetic mapping on one hand and pulse- field gel electrophoresis on the other (Goss and Harris, 1975; Cox *et al.*, 1990). The physical map will provide order and distance estimates for loci the region and will form the basis for isolation of candidate genes and new polymorphic markers

from the region. The physical map is complemented with a genetic map using genotyping data obtained with the polymorphic markers isolated from the region in a set of reference families. The genetic map is essential for refined localization of the HED gene by genetic analysis.

4) study of genetic homogeneity in HED and of the presence of a founder effect in the French Canadian population

HED families from different ethnic origins can be used to study genetic homogeneity in HED and chart the possible origin(s) and dissemination of HED mutation(s). These families might also provide different HED mutations that will facilitate the identification of the HED gene.

HED is disproportionately common in the French Canadian population suggesting the presence is a founder effect for the mutation. Founder effect refers to "the establishment of a new population by a few original founders (in an extreme case, by a single fertilized female) which carry only a small fraction of the total genetic variation of the parental population" (Diamond and Rotter, 1987). A disease mutation carried by a founder will be in linkage disequilibrium (LD) or non-random allelic association with adjacent loci in the resulting population. The French Canadian population was founded by 8000-10,000 immigrants who came mostly from the central and western provinces of France during the period 1608-1769 (Boleda, 1984; Desjardins, 1990). There is strong evidence for a founder effect in many French Canadian genetic diseases such as oculopharyngeal muscular dystrophy (Brais *et al.*, in press), peripheral neuropathy with or without agenesis of the corpus callosum (Casaubon *et al.*, 1996) and pseudo-vitamin D-deficiency rickets (Labuda *et al.*, 1996). We undertook to investigate the presence of

allelic association between HED and polymorphic markers mapped to the candidate region using Fisher's exact test on a 2×2 table. The strength of association will be determined using the formula: $P_{excess} = (P_{affected} - P_{normal})/(1 - P_{normal})$ where P_{normal} and $P_{affected}$ represent the allele frequencies in the general population and disease-carrying chromosomes respectively (Lehesjoki *et al.*, 1993). The affected chromosomes will also be examined for the presence of an ancestral haplotype that most likely carried the founding mutation. This information is important to understand the origin and spread of HED in the French Canadian population. In addition, the presence of a founder effect will allow us to use linkage disequilibrium (LD) mapping to refine the localization of the HED gene as described below.

5) defining the minimum cosegregating HED region by recombination and linkage disequilibrium analyses

The maximum number of HED families are to be collected and new microsatellite markers isolated from YAC and PAC clones in the candidate region for better definition of the smallest chromosomal segment co-inherited with the HED trait. The genetic interval will be defined in two ways: recombination and linkage disequilibrium (LD) analysis.

Recombination analysis involves the identification of recombination events occurring between HED and flanking markers. The haplotypes of all families will be constructed with the markers using the order established by the physical and genetic mapping as described in step 3.

LD mapping relies on the following phenomenon. When a disease mutation is first introduced into a population (by migration or mutation), it resides on a single

chromosome with a unique set of marker alleles forming the ancestral or "founder" haplotype. Consequently, there is complete LD between these marker alleles and the mutation. Through subsequent generations, recombination occurs leading to gradual decay of LD with markers closer to the disease gene showing higher LD than those that are farther away. Since LD reflects recombination events that happened during the history of a population, it increases the number of recombinants and provides a mapping resolution much higher than that of conventional linkage studies. LD depends mainly on the recombination fraction and the number of generations since the introduction of the mutation into the population. Other factors, such as mutation rate, selection pressure, population admixture, genetic drift and growing population size can also affect the pattern of LD (Jorde, 1995).

Various LD mapping methods, simple and model-based, have been developed for fine-scale mapping of disease loci. Simple LD mapping is based solely on measuring the magnitude of the disequilibrium measures to infer the likely position of a disease locus. Model-based methods incorporate the effects of the factors that can affect the LD pattern such as evolutionary history, genetic drift and mutation rate at marker and disease loci (summarized in Xiong and Guo, 1997). For example, the Luria-Delbrück method assumes exponential population growth and estimates the recombination rate by assessing the incidence of the non-associated alleles in the present population (Hästbacka et al., 1992). Terwilliger (1995) developed a likelihood-based method without assumption of the evolutionary sampling process, that uses all alleles at each marker locus and then combines pairwise likelihoods by multiplication for larger number of loci. In contrast, Kaplan *et al.* (1995) proposed a Poisson branching process to model a population and

used simulations of the evolutionary history to provide estimates for the location of the disease locus.

The following chapters present work done as part of this thesis. Chapter 2 will describe the mapping of the HED gene to the pericentromeric region of chromosome 13q by linkage analysis. Chapter 3 will present the most likely exclusion of the candidate genes, connexin 26, connexin 46, fibroblast growth factor 9, zinc finger ZNF198 and α-tubulin TUBA2, for involvement in HED by SSCP analysis. Chapter 4 will depict the construction of a radiation hybrid map of 48 loci including the HED locus and the isolation of eight new polymorphic markers from YAC and PAC clones mapped to the region. Finally, chapter 5 will detail the genetic analysis of HED including genetic homogeneity in a multiethnic group of families, presence of a founder effect in the French Canadian population and fine genetic mapping of the HED gene by linkage and linkage disequilibrium analyses. This work represents an essential step towards the molecular characterization of the defect causing HED. The ultimate goal, identification of the defective gene and characterization of its protein product, will contribute to elucidation of the pathophysiology of HED and will provide insights into the intriguing molecular mechanisms underlying the normal development and healthy maintenance of the skin and its appendages.

1.4. References

Ahmad, W., ul Haque, M. F., Brancolini, V., Tsou, H. C., ul Haque, S., Lam, H., Aita, V. M.; Owen, Jason; deBlaquiere, M., Frank, J., Cserhalmi-Friedman, P. B., Leask, A.,

McGrath, J. A., Peacocke, M., Ahmad, M., Ott, J., and Christiano, A. M. (1998). Alopecia Universalis associated with a mutation in the Human hairless gene. *Science* **279:** 420-424.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215:** 403-410.

Ando, Y., Tanaka, T., Horiguchi, Y., Ikai, K., and Tomono, H. (1988). Hidrotic ectodermal dysplasia: a clinical and ultrastructural observation. *Dermatologica* **176:** 205-211.

Blessing, M., Jorcano, J. L., and Franke, W. W. (1989). Enhancer elements directing cell-type-specific expression of cytokeratin genes and changes of the epithelial cytoskeleton by transfections of hybrid cytokeratin genes. *EMBO J.* **8:** 117-126.

Blessing, M., Zentgraf, H., and Jorcano, J. L. (1987). Differentially expressed bovine cytokeratin genes. Analysis of gene linkage and evolutionary conservation of 5'-upstream sequences. *EMBO J.* **6:** 567-75.

Blumenberg, M., Connolly, D. M., and Freedberg, I. M. (1992). Regulation of keratin gene expression: the role of the nuclear receptors for retinoic acid, thyroid hormone, and vitamin D3. *J. Invest. Dermatol.* **98(6 Suppl):** 42S-49S.

Boleda, M. (1984). Les migrations au Canada sous le régime français (1608-1760). *Cah. Québécois Demogr.* **13:** 23-39.

Bowden, P. E., Hainey, S. D., Parker, G., Jones, D. O., Zimonjic, D., Popescu, N., and Hodgins, M. B. (1998). Characterization and chromosomal localization of human hair-specific keratin genes and comparative expression during the hair growth cycle. *J. Invest. Dermatol.* **110:** 158-64.

Brais, B., Bouchard, J. P., Jomphe, M., Desjardins, B., Dubé, M.-P., Gosselin F., Xie, Y.-G., Tomé, F. M. S., Fardeau, M., Brunet, G., Samson, F., Mathieu, J., Prévost, C., Duranceau, A., Codère, F., Pericak-Vance, M., Bulman, D. E., Tapscott, S., Feasby, T., Munsat, T. L., Pilarski, R., Morgan, K., and Rouleau, G.A. When genetics and history converge: the fine-mapping and North American introduction and diffusion of the French Canadian oculopharyngeal muscular dystrophy mutation. *Am. J. Hum. Genet.* **In Press**.

Casaubon, L. K., Melanson, M., Lopes-Cendes, I., Marineau, C., Andermann, E., Andermann, F., Weissenbach, J., Prévost, C., Bouchard, J. -P., Mathieu, J., and Rouleau, G.A. (1996). The gene responsible for a severe form of peripheral neuropathy and agenesis of the corpus callosum maps to chromosome 15q. *Am. J. Hum. Genet.* **58:** 28-34.

Christiano, A. M., Greenspan, D. S., Hoffman, G. G., Zhang, X., Tamai, Y., Lin, A. N., Dietz, H. C., Hovnanian, A., and Uitto, J. (1993). A missense mutation in type VII collagen in two affected siblings with recessive dystrophic epidermolysis bullosa. *Nature Genet.* **4:** 62-66.

Chuong, C. -M., Chen, H. -M., Jiang, T. -X., and Chia, J. (1991). Adhesion molecules in skin development: morphogenesis of feather and hair. *Ann. N. Y. Acad. Sci.* **642:** 263-280.

Clouston, H. R. (1929). Hereditary ectodermal dystrophy. *Can. Med. Ass. J.* **21:** 18-31.

Clouston, H. R. (1939). The major forms of hereditary ectodermal dysplasia. *Can. Med. Ass. J.* **40:** 1-7.

Couchman, J. R. (1993). Hair follicle proteoglycans. *J. Invest. Dermatol.* **101(1 suppl.):** 60S-64S.

Couchman, J. R., McCarthy, K. J., and Woods, (1991). Proteoglycans and glycoproteins in hair follicle development and cycling. *Ann. NY Acad. Sci.* **642:** 243-51.

Coulombe, P. A., Hutton, M. E., Letai, A., Hebert, A., Paller, A. S., and Fuchs, E. (1991). Point mutations in human keratin 14 genes of epidermolysis bullosa simplex patients: genetic and functional analysis. *Cell* **66:** 1301-1311.

Cox, D. R., Burmeister, M., Price, E. R., Kim, S., and Myers, R. M. (1990). Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. *Science* **250:** 245-250.

Der Kaloustian, V. M., and Kurban, A. K. (1979). In *Genetic diseases of the skin.* Springer-Verlag, pp. 109-114.

Desjardins, B. (1990). Homogénéité ethnique de la population québécoise sous le régime français. *Cah. Québécois Demogr.* **19:** 63-76.

Diamond, J. M., and Rotter, J. I. (1987). Observing the founder effect in human evolution. *Nature* **329:** 105-106.

Escobar, V., Goldblatt, L. I., Bixler, D., and Weaver, D. (1983). Clouston sydrome: an ultrastructural study. *Clin. Genet.* **24:** 140-146.

Freire-Maia, N. (1971). Ectodermal dysplasias. *Hum. Hered.* **21:** 309-312.

Freire-Maia, N., and Pinheiro, M. (1984). In *Ectodermal dysplasias: a clinical and genetic study.* Alan R. Liss, Inc., NewYork, pp. 68-70.

Fuchs, E., Esteves, R. A., and Coulombe, P. A. (1992). Transgenic mice expressing a mutant keratin 10 gene reveal the likely genetic basis for epidermolytic hyperkeratosis. *Proc. Natl. Acad. Sci. USA.* **89:** 6906-6910.

Fuchs, E. (1995). Keratins and the skin. *Annu. Rev. Cell Dev. Biol..***11:** 123-153.

Fuchs, E. (1996a). The cytoskeleton and disease: genetic disorders of intermediate filaments. *Annu. Rev. Genet.* **30:** 197-231.

Fuchs, E. (1996b). Genetic disorders of keratins and their associated proteins. *J. Dermatol. Sci.* **13:** 181-92.

Fuchs, E. (1997).Of mice and men: genetic disorders of the cytoskeleton. *Mol. Biol. Cell* **8:** 189-203.

George, D. I., and Escobar, V. H. (1984). Oral findings of Clouston's syndrome (hidrotic ectodermal dysplasia). *Oral Surg.* **57:** 258-262.

Giraud, F., Mattei, J. F., Rolland, M., Ghiglione, C., Pommier de Santi, P., and Sudan, N. (1977). La dysplasie ectodermique de type Clouston. *Arch. Franç. Péd.* **34:** 982-993.

Gold, R. J. M., and Scriver, C. R. (1972). Properties of hair keratin in an autosomal dominant form of ectodermal dysplasia. *Am. J. Hum. Genet.* **24:** 549-561.

Goliger, J. A., and Paul, D. L. (1994). Expression of gap junction proteins Cx26, Cx31.1, Cx37, and Cx43 in developing and mature rat epidermis. *Dev. Dyn.* **200:** 1-13.

Goss, S. J., and Harris, H. (1975). New method for mapping genes in human chromosomes. *Nature* **255:** 680-4.

Hassed, S. J., Kincannon, J. M., and Arnold, G. J. (1996). Clouston syndrome: an ectodermal dysplasia without significant dental findings. *Am. J. Med. Genet.* **61:** 274-276.

Hästbacka, J., de la Chapelle, A., Kaitila, I., Sistonen, P., Weaver, A., and Lander, E. (1992). Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. *Nature Genet.* **2:** 204-211.

Hazen, P. G., Zamora, I., Bruner, W. E., and Muir, W. A. (1980). Premature cataracts in a family with hidrotic ectodermal dysplasia. *Arch. Dermatol.* **116:** 1385-1387.

Holbrook, K.A. (1988). Structural abnormalities of the epidermally derived appendages in skin from patients with ectodermal dysplasia: insights into developmental errors. *Birth Defects: Original Article Series.* **24**: 15-44.

Holbrook, K. A., and Minami, S. I. (1991). Hair follicle embryogenesis in the human. Characterization of events in vivo and in vitro. *Ann. N. Y. Acad. Sci.* **642:** 167-196.

Holbrook, K. A., Smith, L. T., Kaplan, E. D., Minami, S. A., Hebert, G. P., and Underwood, R. A. (1993). Expression of morphogens during human follicle development in vivo and a model for studying follicle morphogenesis in vitro. *J. Invest. Dermatol.***101(1 Suppl):** 39S-49S.

Ishida-Yamamoto, A., McGrath, J. A., Lam, H., Lizuka, H., Friedman, R. A., and Christiano, A. M. (1997). The molecular pathology of progressive symmetric erythrokeratoderma: a frameshift mutation in the loricrin gene and perturbations in the cornified cell envelope. *Am. J. Hum. Genet.* **61:** 581-589.

Jiang, C. K., Epstein, H. S., Tomic, M., Freedberg, I. M., and Blumenberg, M. (1991). Functional comparison of the upstream regulatory DNA sequences of four human epidermal keratin genes. *J. Invest. Dermatol.* **96:** 162-167.

Jorde, L. B. (1995). Linkage disequilibrium as a gene-mapping tool. *Am. J. Hum. Genet.* **56:** 11-14.

Kaplan, N. L., Hill, W. G., and Weir, B. S. (1995). Likelihood methods for locating disease genes in nonequilibrium populations. *Am. J. Hum. Genet.* **56:** 18-32.

Kere, J., Srivastava, A. K., Montonen, O., Zonana, J., Thomas, N., Ferguson, B., Munoz, F., Morgan, D., Clarke, A., Baybayan, P., Chen, E. Y., Ezer, S., Saariallo-Kere, U., de la Chapelle, A., and Schlessinger, D. (1996). X-linked anhidrotic (hypohidrotic) ectodermal dysplasia is caused by mutation in a novel transmembrane protein. *Nature Genet.* **13:** 409-416.

Kulesh, D. A., Cecena, G., Darmon, Y. M., Vasseur, M., and Oshima, R. G. (1989). Posttranslational regulation of keratins: degradation of mouse and human keratins 18 and 8. *Mol. Cell Biol.* **9:** 1553-1565.

Labuda, M., Labuda, D., Korab-Laskowska, M., Cole, D. E. C., Zietkiewicz, E., Weissenbach, J., Popowska, E., Pronicka, E., Root, A. W., and Glorieux, F. H. (1996). Linkage disequilibrium analysis in young populations: Pseudo-Vitamin D-Deficiency Rickets and the Founder Effect in French Canadians. *Am. J. Hum. Genet.* **59:** 633-643.

Lane, E. B., Rugg, E. L., Navsaria, H., Leigh, I. M., Heagerty, A. H. M., Ishida-Yamamoto, A., and Eady, R. A. J. (1992). A mutation in the conserved helix termination peptide of keratin 5 in hereditary skin blistering. *Nature* **356:** 244-246.

Lathrop, G. M., Lalouel, J. M., Julier, C., and Ott, J. (1984). Strategies for multilocus linkage analysis in humans. *Proc. Nat. Acad. Sci. USA* **81:** 3443-3445.

Lathrop, G. M., Lalouel, J. M., Julier, C., and Ott, J. (1985). Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. *Am. J. Hum. Genet.* **37:** 482-498.

Lee, S. C., Wang, M., McBride, O. W., O'Keefe, E. J., Kim, I. G., and Steinert, P. M. (1993). Human trichohyalin gene is clustered with the genes for other epidermal structural proteins and calcium-binding proteins at chromosomal locus 1q21. *J. Invest. Dermatol.* **100:** 65-68.

Lehesjoki, A.-E., Koskiniemi, M., Norio, R., Tirrito, S., Sistonen, P., Lander, E., and De La Chapelle, E. (1993). Localization of the *EPM1* gene for progressive myoclonus epilepsie on chromosome 21: linkage disequilibrium allows high resolution mapping. *Hum. Mol. Genet.* **2:** 1229-1234.

Lessin, S. R., Huebner, K., Isobe, M., Croce, C. M., and Steinert, P. M. (1988). Chromosomal mapping of human keratin genes: evidence of non-linkage. *J. Invest. Dermatol.* **102:** 171-177.

Mackinnon, P. J., Powell, B. C., Rogers, G. E., Baker, E. G., Mackinnon, R. N., Hyland, V. J., Callen, D. F., and Sutherland, G. R. (1991). An ultra-high sulfur keratin gene of the human hair cuticle is located at 11q13 and cross-hybridizes with sequences at 11p15. *Mammalian Genome* **1:** 53-56.

McGrath, J. A., Gatalica, B., Christiano, A. M., Li, K., Owaribe, K., McMillan, J. R., Eady, R. A., and Uitto, J. (1995). Mutations in the 180-kD bullous pemphigoid antigen (BPAG2), a hemidesmosomal transmembrane collagen (COL17A1). in generalized atrophic benign epidermolysis bullosa. *Nature Genet.* **11:** 83-86.

McGrath, J. A., McMillan, J. R., Shemanko, C. S., Runswick, S. K., Leigh, I. M., Lane, E. B., Garrod D.R., and Eady, R. A. J. (1997). Mutations in the plakophilin 1 gene result in ectodermal dysplasia/skin fragility syndrome. *Nature Genet.* **17:** 240-244.

Mclean, W. H. I., and Lane, B. (1995). Intermediate filaments in disease. *Curr. Opin. Cell Biol.* **7:** 118-125.

Mclean, W. H. I., Rugg, E. L., Lunny, D. P., Morley, S .M., Lane, E. B., Swensson, O., Dopping-Hepenstal, P. J. C., Griffiths, W. A. D., Eady, R. A. J., Higgins, C., Navsaria, H. A., Leigh, I. M., Strachan, T., Kunkeler, L., and Munro, C. S. (1995). Keratin 16 and keratin 17 mutations cause pachyonychia congenita. *Nature Genet.* **9:** 273-278.

McNaughton, P. Z., and Pierson, L. (1976). Hidrotic ectodermal dysplasia in a black mother and daughter. *Arch. Dermatol.* **112:** 1448-1450.

Messenger, A. G. (1993). The control of hair growth: an overview. *J. Invest. Dermatol.* **101(1 Suppl):** 4S-9S.

Messenger, A. G., Elliott, K., Westgate, G. E., and Gibson W. T. (1991). Distribution of extracellular matrix molecules in human hair follicles. *Ann. N. Y. Acad. Sci.* **642:** 253-262.

Monreal, A. W., Zonana, J., and Ferguson, B. (1998). Identification of a new splice form of the eDA1 gene permits detection of nearly all X-linked hypohidrotic ectodermal dysplasia mutations. *Am. J. Hum. Genet.* **63:** 380-389.

Moore, G. P. M., Du Cros, D. L., Isaacs, K., Pisansarakit, P., and Wynn, P. C. (1991). Hair growth induction: roles of growth factors. *Ann. N. Y. Acad. Sci.* **642:** 308-25.

O'Guin, W. M., Schermer, A., Lynch, M., and Sun, T.-T. (1990). Differentiation-specific expression of keratin pairs. In Goldman, R., and Steinert, P. M. (Eds.), *Cellular and molecular biology of intermediate filaments*. Plenum Press, New York, pp. 301-334.

Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. (1989). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5:** 874-879.

Patel, R. A., Bixler, D., and Norins A. L. (1991). Clouston syndrome: a rare autosomal dominant trait with palmoplantar hyperkeratosis and alopecia. *J. Craniofac. Gen. Dev. Biol.* **11:** 176-179.

Peus, D., and Pittelkow, M. R. (1996). Growth factors in hair organ development and the hair growth cycle. *Dermatol. Clin.* **14:** 559-572.

Pierard, G. E., Van Neste, D., and Letot, B. (1979). Hidrotic ectodermal dysplasia. *Dermatologica* **158:** 168-174.

Pinheiro, M., and Freire-Maia, N. (1994). Ectodermal dysplasias: a clinical classification and a causal review. *Am. J. Med. Genet.* **53:** 153-162.

Powell, B. C., and Beltrame, J. S. (1994). Characterization of a hair (wool) keratin intermediate filament gene domain. *J. Invest. Dermatol.* **102:** 171-177.

Powell, B., Crocker, L., and Rogers, G. (1992). Hair follicle differentiation: expression, structure and evolutionary conservation of the hair type II keratin intermediate filament gene family. *Development* **114:** 417-433.

Powell, B. C, Nesci, A., and Rogers, G. E. (1991). Regulation of keratin gene expression in hair follicle differentiation. *Ann. N. Y. Acad. Sci.* **642:** 1-20.

Powell, B. C., and Rogers, G. E. (1990a). Hard keratin IF and associated proteins. In Goldman, R., and Steinert, P. M.(Eds.), *Cellular and molecular biology of intermediate filaments*. Plenum Press, New York, pp. 267-300.

Powell, B. C., and Rogers, G. E. (1990b). Cyclic hair-loss and regrowth in transgenic mice overexpressing an intermediate filament gene. *EMBO J.* **9:** 1485-93.

Powell, B. C., and Rogers, G. E. (1993). Organization and expression of hair follicle genes. *J. Invest. Dermatol.* **101(1 Suppl):** 50S-55S.

Powell, B. C., and Rogers, G. E. (1997). The role of keratin proteins and their genes in the growth, structure and properties of hair. *EXS.* **78:** 59-148.

Pulkkinen, L., Christiano, A. M., Airenne, T., Haakana, H., Tryggvason, K., and Uitto, J. (1994). Mutations in the γ2 chain gene (LAMC2) of kalinin/laminin 5 in the junctional forms of epidermoysis bullosa. *Nature Genet.* **6:** 293- 297.

Pulkkinen, L., Smith, F. J. D., Shimizu, H., Murata, S., Yaoita, H., Hachisuka, H., Nishikawa, T., McLean, W. H. I. and Uitto, J. (1996). Homozygous deletion mutations in the plectin gene (PLEC1) in patients with epidermolysis bullosa simplex associated with late-onset muscular dystrophy. *Hum. Mol. Genet.* **5:** 1539-1546.

Radhakrishma, U., Blouin, J.-L., Mehenni, H., Mehta, T. Y., Sheth, F. J., Sheth, J. J., Solanki, J. V., and Antonarakis S.E. (1997). The gene for autosomal dominant hidrotic ectodermal dysplasia (Clouston syndrome) in a large Indian family maps to the 13q11-q12.1 pericentromeric region. *Am. J. Med. Genet.* **71:** 80-86.

Rajagopalan, K., and Hai Tay, C. (1977). Hidrotic ectodermal dysplasia. Study of a large Chinese pedigree. *Arch. Dermatol.* **113:** 481-485.

Reis, A., Hennies, H. –C., Langbein, L., Digweed, M., Mischke, D., Drechsler, M., Schröck, E., Royer-Pokora, B., Franke, W. W., Sperling, K., and Küster, W. (1994). Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma (EPPK). *Nature Genet.* **6:** 174-179.

Rogers, G. E., Fietz, M. J., and Fratini, A. (1991). Trichohyalin and matrix proteins. *Ann. N. Y. Acad. Sci.* **642:** 64-80.

Rogers, M. A., Langbein, L., Praetzel, S., Moll, I., Krieg, T., Winter, H., and Schweizer, J. (1997). Sequences and differential expression of three novel human type-II hair keratins. *Differentiation* **61:** 187-94.

Rogers, M. A., Nischt, R., Korge, B., Krieg, T., Fink, T. M., Lichter, P., Winter, H., and Schweizer, J. (1995). Sequence data and chromosomal localization of human type I and type II hair keratin genes. *Exp. Cell Res.* **220:** 357-362.

Rosenberg, M., RayChaudhury, A., Shows, T. B., Le Beau, M. M., and Fuchs,E. (1988). A group of type I keratin genes on human chromosome 17: characterization and expression. *Mol. Cell. Biol.* **8:** 722-736.

Rothnagel, J. A., Dominey, A. M., Dempsey, L. D., Longley, M. A., Greenhalgh, D. A., Gagne, T. A., Huber, M., Frenk, E., Hohl, D., and Roop, D. R. (1992). Mutations in the rod domains of keratins 1 and 10 in Epidermolytic Hyperkeratosis. *Science* **257:** 1128-1130.

Russell, L. J., DiGiovanna, J. J., Rogers, G. R., Steinert, P. M., Hashem, N., Compton, J. G., and Bale, S. J. (1995). Mutations in the gene for transglutaminase 1 in autosomal recessive lamellar ichthyosis. *Nature Genet.* **9:** 279-283.

Shapiro, L. J., Yen, P., Pomerantz, D., Martin, E., Rolewic, L., and Mohandas, T. (1989). Molecular studies of deletions at the human steroid sulfatase locus. *Proc. Natl. Acad. Sci. USA* **86:** 8477-8481.

Skalli, O., Chou, Y. -H., and Goldman, R. D. (1992). Intermediate filaments: not so tough after all. *Trends Cell Biol.* **2:** 308-312.

Solomon, L. M., and Keuer, E. J. (1980). The ectodermal dysplasias. Problems of classification and some newer syndromes. *Arch. Dermatol.* **116:** 1295-1299.

Steinert, P. M. (1993). Structure, function, and dynamics of keratin intermediate filaments. *J. Invest. Dermatol.* **100:** 729-734.

Steinert, P. M, and Bale, S. J. (1993). Genetic skin diseases caused by mutations in keratin intermediate filaments. *Trends Genet.* **9:** 280-284.

Steinert, P. M., North, A. C., and Parry D. A. (1994). Structural features of keratin intermediate filaments. *J. Invest. Dermatol.* **103(5 Suppl):** 19S-24S.

Steinert, P. M., and Roop, D. R. (1988). Molecular and cellular biology of intermediate filaments. *Annu. Rev. Biochem.* **57:** 593-625.

Stenn, K. S., Combates, N. J., Eilertsen, K. J., Gordon, J. S., Pardinas J. R., Parimoo, S., and Prouty, S.M. (1996). Hair follicle growth controls. *Dermatol. Clin.* **14:** 543-558.

Takahashi, K., Folmer, J., and Coulombe, P. A. (1994). Increased expression of keratin 16 causes anomalies in cytoarchitecture and keratinization in transgenic mouse skin. *J. Cell Biol.* **127:** 505-520.

Taylor, T.D., Hayflick, S. J., McKinnon, W., Guttmacher, A. E., Hovnanian, A., Litt, M., and Zonana, J. (1998). Confirmation of linkage of Clouston syndrome (hidrotic ectodermal dysplasia) to 13q11-q12.1 with evidence for multiple independent mutations. *J. Invest. Dermatol.* **111:** 83-85.

Terwilliger, J. D. (1995). A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am. J. Hum. Genet.* **56:** 777-787.

Tyner, A. L., and Fuchs, E. (1986). Evidence for posttranscriptional regulation of the keratins expressed during hyperproliferation and malignant transformation in human epidermis. *J. Cell Biol.* **103:** 1945-1955.

Vassar, R., Coulombe, P. A., Degenstein, L., Albers, K., and Fuchs, E. (1991). Mutant keratin expression in transgenic mice causes marked abnormalities resembling a human genetic skin disease. *Cell* **64:** 365-380.

Vidal, F., Aberdam, D., Miquel, C., Christiano, A. M., Pulkkinen, L., Uitto, J., Ortonne, J. P. and Meneguzzi, G. (1995). Integrin beta 4 mutations associated with junctional epidermolysis bullosa with pyloric atresia. *Nature Genet.* **10:** 229-234.

Williams, M., and Fraser, F. C. (1967). Hydrotic ectodermal dysplasia-Clouston's family revisited. *Can. Med. Ass. J.* **96**: 36-38.

Winter, H., Rogers, M. A., Gebhardt, M., Wollina, U., Boxall, L., Chitayat, D., Babul-Hirji, R., Stevens, H. P., Zlotogorski, A., and Schweizer, J. (1997a). A new mutation in the type II hair cortex keratin hHb1 involved in the inherited hair disorder monilethrix. *Hum. Genet.* **101:** 165-169.

Winter, H., Rogers, M. A., Langbein, L., Stevens, H. P., Leigh, I. M., Labrèze, C., Roul, S., Taieb, A., Krieg, T., and Schweizer, J. (1997b). Mutations in the hair cortex keratin hHb6 cause the inherited hair disease monilethrix. *Nature Genet.* **16:** 372-374.

Worobec-Victor, S. M., Bene-Bain, M. A., Shanker, D. B., and Solomon, L. M. (1988). In Schachner, L. A., and Hansen, R. C. (eds.), *Pediatric Dermatology.* Chuchill Livingstone, Vol.I, p. 328.

Xiong, M., and Guo, S. W. (1997). Fine-scale genetic mapping based on linkage disequilibrium: theory and applications. *Am. J. Hum. Genet.* **60:** 1513-1531.

Yu, J., Yu, D. -W., Checkla, D. M., Freedberg, I. M., and Bertolino, A. P. (1993). Human hair keratins. *J. Invest. Dermatol.* **101(1 Suppl):** 56S-59S.

CHAPTER 2

The gene responsible for Clouston hidrotic ectodermal dysplasia maps to the pericentromeric region of chromosome 13q

This was the first step towards the identification of the HED gene, which is mapping the gene to a chromosomal region by linkage analysis.

The presented work was published as:

Kibar, Z., Der Kaloustian, V. M., Brais, B., Hani, V., Fraser, F. C., and Rouleau, G. A. (1996). The gene responsible for Clouston hidrotic ectodermal dysplasia maps to the pericentromeric region of chromosome 13q. *Hum. Mol. Genet.* **5:** 543-547.

In this chapter, references were reformatted and listed in alphabetical sequence by author's last name to be consistent with other chapters in the thesis.

Contribution by co-authors:

Kibar: DNA extraction and genotyping; linkage analysis

Der Kaloustian, Hani and Fraser: collection of HED family material

Brais: helpful suggestions for the linkage analysis

Rouleau: lab. supervisor

ABSTRACT

Hidrotic ectodermal dysplasia (HED), Clouston type, is an autosomal dominant skin disorder which is most common in the French Canadian population and is characterized by hair defects, nail dystrophy and palmoplantar hyperkeratosis. Biophysical and biochemical studies conducted in HED suggested a molecular abnormality of keratins. We tested eight French Canadian families segregating HED for linkage to microsatellite markers flanking the known keratin genes and were able to exclude linkage to these loci. Therefore, a genome-wide search for the HED gene was initiated. The first lod score above 3.00 was obtained with the marker D13S175 located in the pericentromeric region of chromosome 13q ($Z_{max} = 8.12$ at zero recombination). The cumulative lod scores were above 3.00 for six other markers in the region. A multipoint linkage analysis using the markers D13S175, D13S141 and D13S143 gave a maximum lod score of 11.12 at D13S141 with the one-lod-unit support interval spanning a 12.7 cM region which includes D13S175 and D13S141. Haplotype analysis allowed us to establish D13S143 as the telomeric flanking marker for the HED candidate region.

INTRODUCTION

Hidrotic ectodermal dysplasia (HED), or Clouston syndrome, is a rare autosomal dominant genodermatosis transmitted with complete penetrance and variable clinical expressivity (Williams and Fraser, 1967). It was first described by Nicolle and Hallipré in a French Canadian family in 1895 (George and Escobar, 1984). The hereditary nature of HED was established only in 1929 when Clouston described the disorder in five generations of a large French Canadian family living in Québec (Clouston, 1929). To

date, most of the families segregating this condition are of either French or of French Canadian origin, suggesting a common ancestor (Escobar *et al.*, 1983). Occurrence in families of other ethnic origins such as Chinese (Rajagopolan and Hai Tay, 1977), Japanese (Ando *et al.*, 1988), British (Patel *et al.*, 1991) and African-American (McNaughton and Pierson, 1976) has also been reported. There is no obvious genetic origin common to all cases.

The main features of HED are dystrophy of the nails that tend to be hypoplastic and deformed with increased susceptibility to paronychial infections, defects of the hair that range from brittleness and slow growth rate to total alopecia and moderate to severe palmoplantar hyperkeratosis with reduced keratinocyte desquamation (Der Kaloustian and Kurban, 1979; Ando *et al.*, 1988). Variable and infrequent features reported with HED include ocular abnormalities, sensorineural deafness, hyperpigmentation of the skin, mental retardation and skeletal anomalies such as polydactyly and syndactyly (Worobec-Victor *et al.*, 1988). In contrast to the X-linked anhidrotic type of ectodermal dysplasia, facial appearance, tooth growth and sweating capacity are normal in HED (Der Kaloustian and Kurban, 1979).

Biophysical and biochemical studies conducted in HED suggest depletion of a component of the high-sulfur portion of the matrix proteins related to a disruption of disulfide bond formation in the remaining keratin. The hair is thin with abnormal swelling capacity, reduced tensile strength, reduced birefringence under polarized light, looser and disorganized structure by light microscopy and decreased levels of cysteine and disulfide bonds (Gold and Scriver, 1972). The diminished cysteine content was also shown in a biochemical study of keratin in the nails of HED patients (Giraud *et al.*, 1977).

Ultrastructural findings reveal disorganization of hair fibrils with desquamation of the cuticle that could be due to the reduction of disulfide bonds (Escobar *et al.*, 1983). All these abnormal changes are suggestive of a molecular defect involving the keratin proteins. Point mutations in keratin genes have been described in many ectodermal dysplasias, such as epidermolysis bullosa simplex (Hovnanian et al., 1993), epidermolytic hyperkeratosis (Rothnagel *et al.*, 1992), and pachyonychia congenita (Mclean *et al.*, 1995), stressing their importance in maintaining the structural integrity of and providing mechanical strength to the epithelial cells (Compton, 1994).

In the present study, we report the chromosomal localization of the gene responsible for HED to the pericentromeric region of chromosome 13q using linkage analysis of nine microsatellite markers in eight French Canadian families.

RESULTS

Two-point Linkage analysis

Eight French Canadian families segregating HED were studied (figure 2.1.). They included 45 affected among a total of 76 participants. First, we sought linkage of HED to the chromosomal regions harboring the candidate keratin genes. Epidermal soft keratin genes have been found in clusters on two chromosomes: the type I acidic cluster on chromosome 17 and the type II basic cluster on chromosome 12 (Lessin *et al.*, 1988; Rosenberg *et al.*, 1988). Many of the hair keratin genes have been isolated and characterized but not mapped yet (Rogers and Powell, 1993; Bowden *et al.*, 1994); however, there is evidence from linkage data in mice that they are associated with the epidermal keratin genes (Compton *et al.*, 1991). Other chromosomes that have been

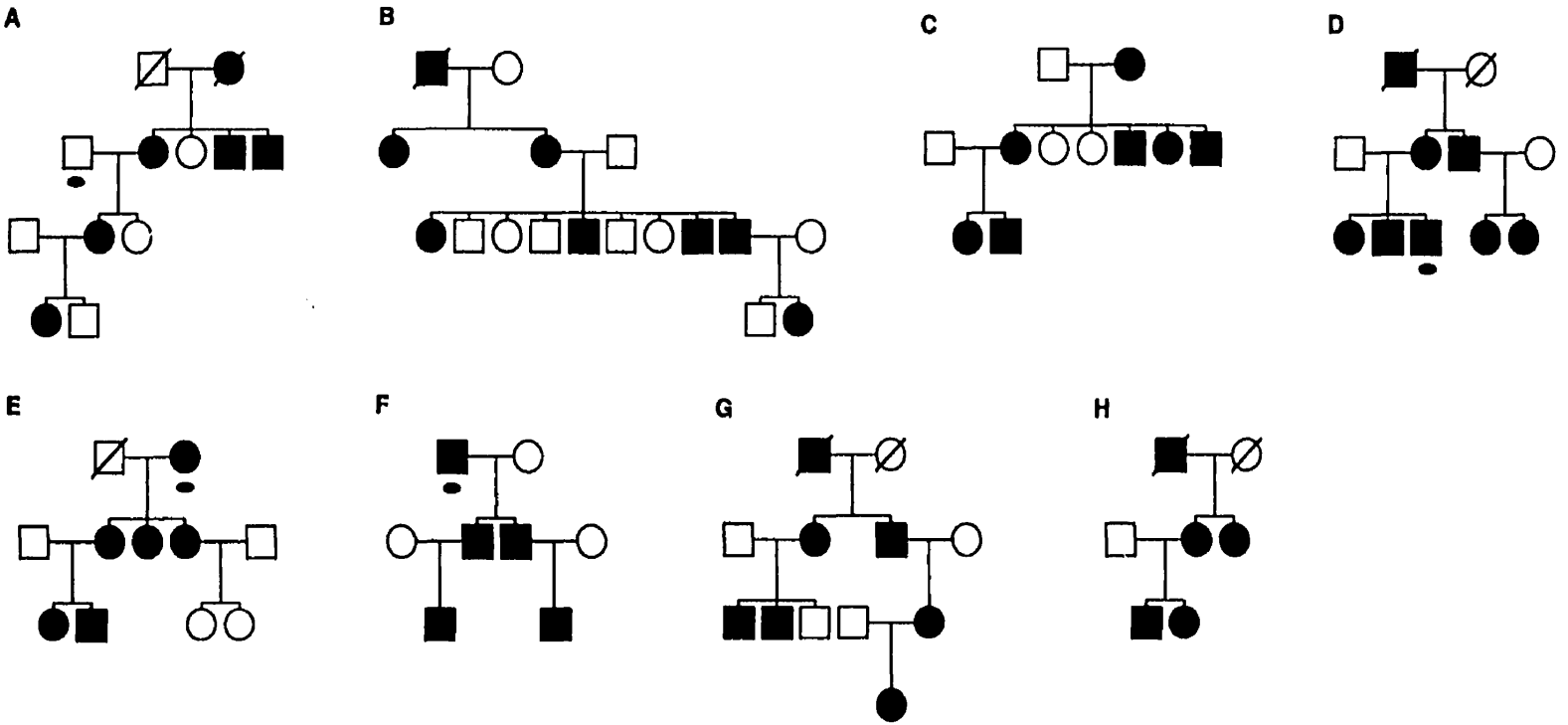

Figure 2. 1. Pedigrees of the HED families used in this study. DNA is available from all live members shown except those indicated by black ovals.

suggested as the sites of more distantly related keratin genes, by in situ hybridization techniques, include chromosomes 14, 15 and 21 (Rosenberg *et al.*, 1988). Genes encoding the cuticle's ultrahigh-sulfur proteins have been localized to two clusters on chromosome 11 (Mackinnon *et al.*, 1991). We started by typing the HED pedigrees with polymorphic microsatellite markers localized to these chromosomal regions and we found no evidence of linkage to any of them. Consequently, we initiated a general genome search using polymorphic markers spaced on average 20 cM apart. After excluding approximately 10% of the genome by typing 27 polymorphic markers, we found linkage to the pericentromeric marker D13S175 with a maximum combined two-point lod score of 8.12 at $\theta = 0$. Family B, by itself, gave a maximum lod score of 3.52 at $\theta = 0$ (table 2.1.).

To confirm the initial linkage and to define the HED locus further, pairwise linkage analyses were performed with eight additional markers mapped to the pericentromeric region of chromosome 13. Five of these microsatellites were obtained from Genethon: D13S1316, D13S1236, D13S1275, D13S292 and D13S283; the other three were physically mapped by somatic cell hybridization: D13S141 and D13S143 to 13q11-q12.1 (Pertrukhin *et al.*, 1993) and D13S115 to 13pter-q12.1 (Hudson *et al.*, 1992). Table 2.1. summarizes the two-point lod scores for all nine markers at various recombination fractions. Lod scores >3.00 were obtained for all of the markers except D13S1236 and D13S283. D13S1236 was not informative in this set of families and D13S283 which maps 5 cM telomeric to D13S175 (Gyapay *et al.*, 1994) excluded linkage to HED up to $\theta = 0.04$ ($Z_{max} = -2.64$). Three markers showed no recombination with a maximum combined lod score >3 at $\theta = 0$: D13S1316, D13S175

Table 2. 1. Two-point lod scores between HED and nine markers mapped to the pericentromeric region of chromosome 13q at various recombination fractions

Marker	Family	Lod score (Z) at recombination fraction θ							Z_{max}	θ
		0.00	0.05	0.10	0.15	0.20	0.25	0.30		
D13S1316	B	0.08	0.07	0.05	0.04	0.03	0.02	0.01		
	total	3.83	3.41	2.98	2.55	2.12	1.68	1.24	3.83	0.00
D13S1236	B	0.12	0.10	0.08	0.07	0.05	0.03	0.02		
	total	$-\infty$	2.22	2.46	2.32	2.03	1.67	1.27	2.46	0.10
D13S175	B	3.52	3.24	2.95	2.65	2.33	2.00	1.65		
	total	8.12	7.39	6.61	5.80	4.95	4.10	3.23	8.12	0.00
D13S1275	B	$-\infty$	1.37	1.45	1.39	1.27	1.12	0.93		
	total	$-\infty$	4.14	4.41	4.15	3.68	3.08	2.41	4.42	0.09
D13S292	B	$-\infty$	1.74	1.79	1.71	1.56	1.37	1.15		
	total	$-\infty$	3.01	3.67	3.67	3.40	2.97	2.43	3.71	0.12
D13S283[a]	B	$-\infty$	-1.16	-0.42	-0.07	0.12	0.21	0.25		
	total	$-\infty$	-1.92	-0.01	0.76	1.06	1.11	1.00	1.11	0.22
D13S141	B	1.21	1.10	0.99	0.88	0.76	0.63	0.51		
	total	4.00	3.54	3.08	2.61	2.16	1.71	1.28	4.00	0.00
D13S143	B	$-\infty$	2.00	2.03	1.93	1.75	1.54	1.29		
	total	$-\infty$	7.14	6.83	6.18	5.37	4.45	3.48	7.14	0.05
D13S115	B	$-\infty$	1.45	1.53	1.47	1.35	1.19	1.00		
	total	$-\infty$	4.00	4.34	4.15	3.72	3.16	2.52	4.34	0.10

[a] Z_{max}=-2.64 at recombination fraction of 0.04

and D13S141. D13S175 gave the highest two-point lod score in this study of 8.12 at θ = 0.

Haplotype and multipoint linkage analyses

In order to define the smallest interval containing the HED locus, we analyzed the eight families for recombination events between marker loci and HED. Since no single chromosome 13 map contains all the markers studied, we first constructed the haplotypes using only the ordered genethon markers as follows: D13S1316-D13S1236-D13S175-D13S1275-D13S292-D13S283 (Genethon). Using this order, we detected 3 double recombinants in 3 families (data not shown). Our data suggested an order consistent with physical data for a 6 Mb YAC contig established in the 13q pericentromeric region spanning the region between D13S175 and D13S221, establishing the following order for the markers we typed: centromere-D13S175-D13S141-D13S143-D13S115-D13S1236-D13S292-D13S283-telomere (Guilford *et al.*, 1995). D13S1316 and D13S1275 were not included in the haplotype analysis because their position has yet to be determined. We identified eight informative recombination events in five families (data not shown). Figure 2.2. shows the two critical crossovers that would place the HED locus centromeric to D13S143 (individuals III-3 in family A and III-3 in family B). In addition, we found that all 45 affected individuals share a common haplotype (2 2 3) for the markers D13S175, D13S141 and D13S143.

To estimate the most likely position of the HED gene, we performed multipoint linkage analysis using the genetic map: D13S175-2 cM-D13S141-6 cM-D13S143. The recombination fractions between these markers were determined from pairwise linkage

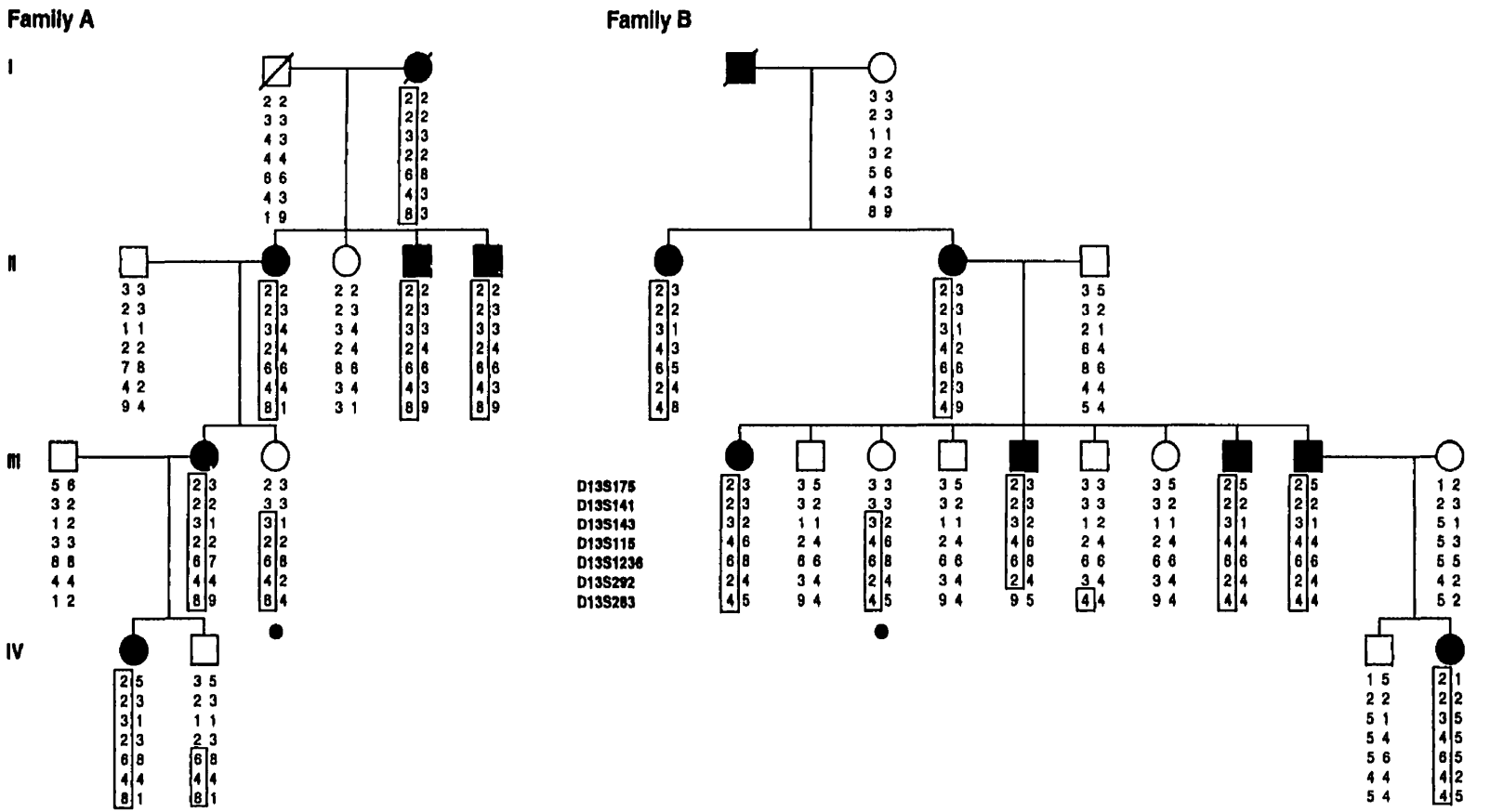

Figure 2. 2. Pedigrees of two HED families with their haplotypes showing two recombinants, individuals III-3 in family A and III-3 in family B, that would place the HED locus proximal to D13S143. The phasing of haplotypes was inferred from the available genotypic data. Haplotypes that most likely contain the mutant HED gene are boxed. The order of the markers used in this analysis was obtained from a YAC contig established in the region (Guilford *et al.*, 1995).

analysis in the eight HED kindreds (data not shown). The maximum multipoint lod score attained was 11.12 at D13S141. However, if we drop 1 lod unit below the maximum, the support interval would span a 12.7 cM region, 6.3 cM centromeric to D13S175 and 4.4 cM telomeric to D13S141 (figure 2.3.).

DISCUSSION

By using linkage analysis in eight French Canadian families segregating HED, we were able to map the responsible gene to the pericentromeric region of chromosome 13q. Strong evidence for linkage to HED was obtained with seven polymorphic markers mapped to this region of which D13S175 generated the highest cumulative two-point lod score ($Z_{max} = 8.12$ at $\theta = 0$). Data from haplotype analysis showed that D13S143 is the telomeric flanking marker for the candidate region of the HED gene. In the absence of more centromeric markers, we could not define a proximal limit to this region. However, multipoint linkage analysis places the HED locus at the most 6.3 cM centromeric to D13S175.

Since the biochemical defect in HED is believed to involve keratins, exclusion of linkage to all chromosomes known to harbor the keratin genes i.e. chromosomes 11, 12, 14, 15, 17 and 21 was unexpected. However, not all keratin genes have been mapped and the HED locus could still correspond to one of an unlocalized keratin. Another hypothesis has been proposed to explain all types of ectodermal dysplasias, including HED, where the disease is caused by a disturbed mesoderm-ectoderm interaction during morphogenesis of the ectodermal tissues (Holbrook, 1988; Pierard *et al.*, 1979). Given this hypothesis, it may be significant that two forms of non syndromic neurosensory

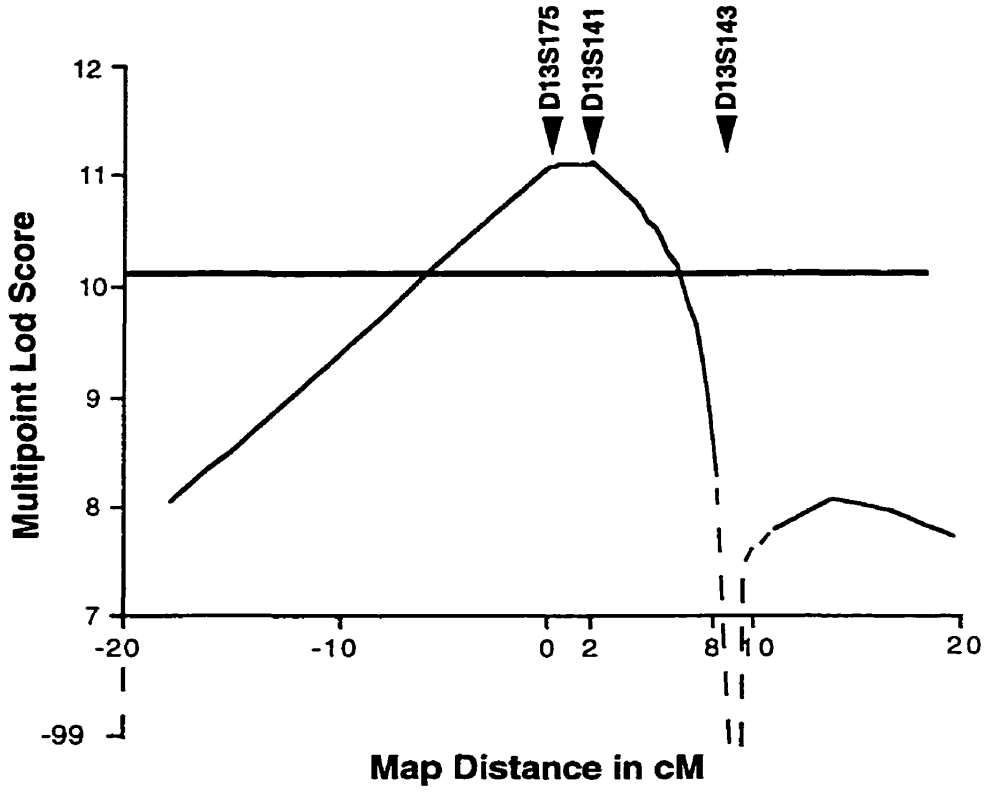

Figure 2. 3. Multipoint linkage analysis of HED with the markers D13S175, D13S141 and D13S143. The solid line represents the 1 lod-unit support interval for the placement of the HED locus. The genetic distance between markers was determined from two-point linkage analysis in the eight HED families. The map distance was calculated using the Haldane mapping function.

deafness, recessive DFNB1 (Guilford *et al.*, 1994) and dominant DFNA3 (Chaïb *et al.*, 1994), also map to the region containing the HED locus and show linkage to D13S175, D13S143 and D13S115. These two diseases result from an endocochlear defect and the responsible genes may code for one of the proteins involved in cochlea structure and function. Because cochlea development is influenced and controlled by a series of inductive ectoderm-mesoderm reciprocal interactions (Van de Water *et al.*, 1988) and because hearing loss has been reported in a few cases of HED (Der Kaloustian and Kurban, 1979), it is possible that these 3 diseases are caused by different mutations in the same gene or in related genes found in a cluster.

The candidate region for the HED gene contains three expressed sequence tags (ESTs) (D13S183E, D13S502E and D13S837E) and an α-tubulin gene (TUBA2) (Guilford *et al.*, 1995). These ESTs will be further analyzed to establish their possible role as candidate genes for HED. TUBA2 is unlikely to be involved in this disease since its expression is most probably testis-specific (Guilford *et al.*, 1995). We will continue to reduce the candidate region by typing new markers and collecting additional HED families. Furthermore, since a founder effect has been suggested by the high prevalence of HED in the French Canadian population, we may be able to use linkage disequilibrium mapping to narrow down the HED candidate region.

MATERIALS AND METHODS

HED families

Diagnosis of HED was established when individuals presented with various degrees of hair abnormalities, nail defects and hyperkeratosis of palms and soles. Eight

French-Canadian families participated in this study consisting of 76 individuals that included 45 affected (figure 2.1., page 39).

DNA analysis and genotyping

From each consenting member, 20 ml blood were obtained: 10 ml were used to establish Epstein-Barr virus transformed lymphoblastoid cell lines and the other 10 ml were used for direct DNA extraction using standard procedures (Sambrook *et al.*, 1989). PCR amplifications for the microsatellite markers were carried out in a total volume of 14 µl containing: 40 ng of genomic DNA; 200 µM of dCTP, dGTP and dTTP; 25 µM dATP, 125 ng of each primer; 1.5 µCi[^{35}S]dATP, 0.5 U Taq DNA polymerase (Perkin Elmer) and 1.25 µl Taq buffer with 15 mM $MgCl_2$ (Perkin Elmer). For the Genethon markers, PCR conditions consisted of 35 cycles with 45 s at 94°C, 30s at 57°C and 45 s at 72°C for each cycle. For D13S141, D13S143 and D13S115, PCR conditions were as previously published (Hudson *et al.*, 1992; Pertrukhin *et al.*, 1993). For D13S143, the primers' sequences were as described for marker ca016(A) (Pertrukhin *et al.*, 1993). The PCR products were run on denaturing 6% polyacrylamide gels in 1×TBE for 2-3 hrs. After drying, the gels were exposed to X-ray films for 2-3 days at room temperature. An M13 sequencing ladder was used to determine the alleles' sizes for the genotypes. For D13S175, allele 2 is 103 bp; for D13S141, allele 2 is 125 bp and for D13S143 allele 3 is 128 bp.

Linkage analysis

Two-point and multipoint linkage analyses were performed using the MLINK and the LINKMAP programs of the LINKAGE 5.1 package (Lathrop *et al.*, 1984; Lathrop *et al.*, 1985). The mode of inheritance was considered to be autosomal dominant with complete penetrance and an estimated gene frequency of 0.00001. Recombination fractions were assumed to be equal in females and males. The allele frequencies in the French Canadian population for all markers were obtained from the genotypes of 18 unaffected spouses participating in this study.

ACKNOWLEDGMENTS

We would like to thank the families who made this study possible. We thank Dr. Jean Morissette, Dr. Iscia Lopes Cendes and Marie-Pierre Dubé for their invaluable help. This study was supported by the Medical Research Council of Canada, the Fonds de recherche en santé du Québec and the Musto Foundation.

REFERENCES

Ando, Y., Tanaka, T., Horiguchi, Y., Ikai, K., and Tomono, H. (1988). Hidrotic ectodermal dysplasia. *Dermatologica* **176:** 205-211.

Bowden, P. E., Hainey, S., Parker, G., and Hodgins, M. B. (1994). Sequence and expression of human keratin genes. *J. Dermatol. Sci.* 7: S152-S163.

Chaïb, H., Lina-Granade, G., Guilford, P., Plauchu, H., Levilliers, J., Morgon, A., and Petit, C. (1994). A gene responsible for a dominant form of neurosensory non-

syndromic deafness maps to the NSRD1 recessive deafness gene interval. *Hum. Mol. Genet.* **3**: 2219-2222.

Clouston, H. R. (1929) Hereditary ectodermal dystrophy. *Can. Med. Ass. J.*, **21**, 18-31.

Compton, J. G. (1994). Epidermal disease: faulty keratin filaments take their toll. *Nature Genet.* **6**: 6-7.

Compton, J. G., Ferrara, D. M., Yu, D., Recca, V., Freedberg, I. M., and Bertolino, A. P. (1991). Chromosomal localization of mouse hair keratin genes. *Ann. N. Y. Acad. Sci.* **642**: 32-43.

Der Kaloustian, V. M., and Kurban, A. K. (1979). *Genetic diseases of the skin.* Springer-Verlag, pp. 109-114.

Escobar, V., Goldblatt, L. I., Bixler, D., and Weaver, D. (1983). Clouston sydrome: an ultrastructural study. *Clinical Genetics* **24**: 140-146.

George, D. I., and Escobar, V. H. (1984). Oral findings of Clouston's syndrome (hidrotic ectodermal dysplasia). *Oral Surg.* **57**: 258-262.

Giraud, F., Mattei, J. F., Rolland, M., Ghiglione, C., Pommier de Santi, P., and Sudan, N. (1977). La dysplasie ectodermique de type Clouston. *Arch. Franç. Péd.* **34:** 982-993.

Gold, R. J. M., and Scriver, C. R. (1972). Properties of hair keratin in an autosomal dominant form of ectodermal dysplasia. *Am. J. Hum. Genet.* **24:** 549-561.

Guilford, P., Ben Arab, S., Blanchard, S., Levilliers, J., Weissenbach, J., Belkahia, A., and Petit, C. (1994). A non-syndromic form of neurosensory, recessive deafness maps to the pericentromeric region of chromosome 13q. *Nature Genet.* **6:** 24-28.

Guilford, P., Dodé, C., Crozet, F., Blanchard, S., Chaïb, H., Levilliers, J., Levi-Acobas, F., Weil, D., Weissenbach, J., Cohen, D., Le Paslier, D., Kaplan, J. -C., and Petit, C.

(1995). A YAC contig and an EST map in the pericentromeric region of chromosome 13 surrounding the loci for neurosensory nonsydromic deafness (DFNB1 and DFNA3) and limb-girdle muscular dystrophy type 2C (LGMD2C). *Genomics* **29:** 163-169.

Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M., and Weissenbach, J. (1994). The 1993-1994 généthon human genetic linkage map. *Nature Genet.* **7:** 246-339.

Holbrook, K. A. (1988). Structural abnormalities of the epidermally derived appendages in skin from patients with ectodermal dysplasia: insights into developmental errors. *Birth Defects Original Article Series.* **24:** 15-44.

Hovnanian, A., Pollack, E., Hilal, L., Rochat, A., Prost, C., Barrandon, Y., and Goossens, M. (1993). A missense mutation in the rod domain of keratin 14 associated with recessive epidermolysis bullosa simplex. *Nature Genet.* **3:** 327-332.

Hudson, T. J., Engelstein, M., Lee, M. K., Ho, E. C., Rubenfield, M. J., Adams, C. P., Housman, D. E., and Dracopoli, N. C. (1992). Isolation and chromosomal assignment of 100 highly informative human simple sequence repeat polymorphisms. *Genomics* **13:** 622-629.

Lathrop, G. M., Lalouel, J. M., Julier, C., and Ott, J. (1984). Strategies for multilocus linkage analysis in humans. *Proc. Nat. Acad. Sci. USA* **81:** 3443-3445.

Lathrop, G. M., Lalouel, J. M., Julier, C., and Ott, J. (1985). Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. *Am. J. Hum. Genet.* **37:** 482-498.

Lessin, S. R., Huebner, K., Isobe, M., Croce, C. M., and Steinert, P. M. (1988). Chromosomal mapping of human keratin genes: evidence of non-linkage. *J. Invest. Dermatol.* **102:** 171-177.

Mackinnon, P. J., Powell, B. C., Rogers, G. E., Baker, E. G., Mackinnon, R. N., Hyland, V. J., Callen, D. F., and Sutherland, G. R. (1991). An ultra-high sulfur keratin gene of the human hair cuticle is located at 11q13 and cross-hybridizes with sequences at 11p15. *Mammalian Genome* **1:** 53-56.

Mclean, W. H. I., Rugg, E. L., Lunny, D. P., Morley, S. M., Lane, E. B., Swensson, O., Dopping-Hepenstal, P. J. C., Griffiths, W. A. D., Eady, R. A. J., Higgins, C., Navsaria, H. A., Leigh, I. M., Strachan, T., Kunkeler, L., and Munro, C. S. (1995). Keratin 16 and keratin 17 mutations cause pachyonychia congenita. *Nature Genet.* **9:** 273-278.

McNaughton, P. Z., and Pierson, L. (1976). Hidrotic ectodermal dysplasia in a black mother and daughter. *Arch. Dermatol.* **112:** 1448-1450.

Patel, R. A, Bixler, D., and Norins, A. L. (1991). Clouston syndrome: a rare autosomal dominant trait with palmoplantar hyperkeratosis and alopecia. *J. Craniofac. Gen. Dev. Biol.* **11:** 176-179.

Pertrukhin, K. E., Speer, M. C., Cayanis, E., Bonaldo, M. D. F., Tantravahi, U., Soares, M. B., Fischer, S. G., Warburton, D., Gilliam, T. C., and Ott, G. (1993). A microsattelite genetic linkage map of human chromosome 13. *Genomics* **15:** 76-85.

Pierard, G. E., Van Neste, D., and Letot, B. (1979). Hidrotic ectodermal dysplasia. *Dermatologica* **158:** 168-174.

Rajagopalan, K., and Hai Tai, C. (1977). Hidrotic ectodermal dysplasia. *Arch. Dermatol.* **113:** 481-485.

Rogers, G. E., and Powell, B. C. (1993). Organization and expression of hair follicles. *J. Invest. Dermatol.* **101:** 50S-55S.

Rosenberg, M., RayChaudhury, A., Shows, T. B., Le Beau, M. M., and Fuchs, E. (1988). A group of type I keratin genes on human chromosome 17: characterization and expression. *Mol. Cell. Biol.* **8:** 722-736.

Rothnagel, J. A., Dominey, A. M., Dempsey, L. D., Longley, M. A., Greenhalgh, D. A., Gagne, T. A., Huber, M., Frenk, E., Hohl, D., and Roop, D. R. (1992). Mutations in the rod domains of keratins 1 and 10 in Epidermolytic Hyperkeratosis. *Science* **257:** 1128-1130.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A laboratory Manual.* Cold Spring Harbor Laboratory Press. Vol.2, pp. 9.16-9.19.

Van de Water, T. R., Noden, D. M., and Maderson, P. F. A. (1988). In Alberti,P.W. and Ruben,R.J. (eds). *Otologic Medicine and Surgery.* Churchill Livingstone, Vol.1 pp.3-27.

Williams, M., and Fraser,F.C. (1967). Hydrotic ectodermal dysplasia-Clouston's family revisited. *Can. Med. Ass. J.*, **96:** 36-38.

Worobec-Victor, S. M., Bene-Bain, M. A., Shanker, D. B., and Solomon, L. M. (1988). In Schachner,L.A. and Hansen,R.C. (eds.), *Pediatric Dermatology.* Chuchill Livingstone, Vol.I, p. 328.

CHAPTER 3

Connexin 26, connexin 46, fibroblast growth factor 9, zinc-finger *ZNF198* and α tubulin *TUBA2* are most likely not defective in Clouston hidrotic ectodermal dysplasia

These five genes map to the HED candidate region and hence were tested for involvement in HED

ABSTRACT

Using linkage analysis in eight French Canadian families, we have mapped the gene responsible for Clouston hidrotic ectodermal dysplasia (HED) to the pericentromeric region of chromosome 13q flanked distally by the marker D13S143. The candidate region contains five known genes: connexin 26 (*Cx26*), connexin 46 (*Cx46*), fibroblast growth factor 9 (*FGF-9*), zinc-finger *ZNF198* and α tubulin *TUBA2*. We have cloned the human Cx46 and tested this as well as the other candidate genes for the presence of pathogenic mutations in HED patients using PCR-SSCP analysis. No mutation specific to HED was found in any of the genes tested suggesting that they most likely are not defective in this disease. While this study was in process, pathogenic mutations in *Cx26* and *Cx46* have been identified in patients with hereditary non-syndromic sensorineural deafness (Kelsell *et al.*, 1997) and autosomal dominant "zonular pulverulent" cataract (Mackay *et al.*, 1999) respectively.

INTRODUCTION

Linkage analysis in eight French Canadian families allowed us to map the gene responsible for HED to the pericentromeric region of chromosome 13q. The candidate region spans at least 15 cM and extends from the centromere to D13S143 at 13q11-q12.1 (Kibar *et al.*, 1996). Five genes map to this region and hence were considered as candidates for involvement in HED. They consist of: *Cx26*, *Cx46*, *FGF-9* and, a novel zinc-finger gene, *ZNF198*, mapped to 13q11-q12 by in situ hybridization (Mignon *et al.*, 1996; Mattei *et al.*, 1995; Reiter *et al.*, 1998, respectively) and *TUBA2* mapped to an EST and YAC contig established in the region (Guilford *et al.*, 1995). In the absence of a clear

pathogenetic mechanism in HED, a broad range of candidate genes needs to be considered.

Connexins constitute a multigene family of at least 13 proteins in mammals, which oligomerize to form intercellular channels directly linking the cytoplasm of two adjacent cells. These channels allow the reciprocal exchange of ions and small molecules contributing significantly to a number of cellular processes including homeostasis, synchronization of excitable cells' activities and regulation of growth, differentiation and developmental signaling. The highly related connexins are named based on their molecular weights (in kDa) and can be categorized into two classes, α and β, according to sequence similarities. The expression pattern of the various connexin isoforms is tissue-specific and developmentally regulated with most tissues expressing more than one connexin type in an overlapping manner (Kumar and Gilula, 1992; Bruzonne *et al.*, 1996; Goodenough *et al.*, 1996). The skin utilizes at least seven connexins (Cx-26, -31, -31.1, -37, -40, -43 and -45) in a complex and overlapping pattern of expression that varies with the stage of epidermal and follicular differentiation (Risek *et al.*, 1992; Kamibayashi *et al.*, 1993; Butterweck *et al.*, 1994; Goliger *et al.*, 1994; Salomon *et al.*, 1994).

The human *Cx26* or gap junction β2 gene was cloned by subtractive hybridization selecting for mRNAs expressed in normal human mammary epithelium cells but not in mammary tumor cell lines. The cDNA encodes a putative protein of 226 amino acids with conserved structural characteristics. *Cx26* is down regulated in tumor tissue and thus considered as a growth suppressor gene (Lee *et al.*, 1992). Using a combination of immunohistological studies, western blot and northern blot analyses, Cx26 was shown to be utilized during skin and hair development in rodents (Risek *et al.*, 1992; Kamibayashi

et al., 1993; Butterweck *et al.*, 1994; Goliger *et al.*, 1994) and abundantly expressed in hair follicles in humans (Salomon *et al.*, 1994). *Cx26* is widely expressed in other tissues as well including liver, mammary gland and pancreas (Paul *et al.*, 1995; Bruzzone *et al.*, 1996). Knockout mice deficient in Cx26 die in utero 10-11 days *post coitum* due to impaired glucose in the labyrinth region of the placenta where Cx26 is normally expressed (Nicholson and Bruzzone, 1997). While our study was in process, pathogenic mutations in *Cx26* have been identified in hereditary non-syndromic sensorineural deafness (Kelsell *et al.*, 1997; Kelley *et al.*, 1998).

The *Cx46* or gap junction α3 gene was first isolated from a rat lens cDNA library. The 2.8 Kb mRNA is highly expressed in the lens and to a lesser degree in the heart and kidney (Paul *et al.*, 1991). Mice deficient in Cx46 develop nuclear cataract associated with proteolysis of the crystallins (Gong *et al*, 1997). We have aided in the cloning of the coding sequence of human *Cx46* (as described below) and mutations in this sequence have been found in patients with dominant "zonular pulverulent" cataract (Mackay *et al.*, in 1999).

Fibroblast growth factors (FGFs) are regulatory polypeptides that act through four different tyrosine kinase receptors to mediate cell proliferation, differentiation and function. They are involved in various biological processes including normal development, angiogenesis and tumorigenesis (Galzie *et al.*, 1997). Members of this family are expressed at many stages and in different tissues during embryogenesis and adult life with each member displaying a distinct pattern of expression (Wilkie *et al.*, 1995). FGF signaling has been implicated in the control of mesoderm-ectoderm interaction during the development of the skin and its appendages (du Cros, 1993; Peus

and Pittelkow, 1996). Mice deficient in FGF-5 have a normal follicle but the hair is abnormally long indicating a role for this factor in the regulation of the hair growth cycle (Hébert *et al.*, 1994). FGF-9 or glia-activating factor (GAF) constitutes the ninth member of the FGF family. The protein was purified from the culture supernatant of a human glioma cell line and found to promote glial cell proliferation. Using oligonucleotide probes designed from partial sequence of purified GAF, the FGF-9 cDNA was cloned and found to encode a polypeptide of 208 amino acids. Northern analysis detected expression of the gene in the brain and kidney of the adult rat (Miyamoto *et al.*, 1993). Expression studies of various growth factors demonstrated absence of FGF-9 mRNA in the murine and human anagen hair follicles (Rosenquist and Martin, 1996; Mitsui *et al.*, 1997). Recently, FGF-9 was demonstrated to participate significantly in the signaling events that take place during tooth morphogenesis (Kettumen and Thesleff, 1998).

ZNF198 (also called *RAMP* or *FIM*) is a novel gene isolated in efforts to identify genes involved in t(8;13)(p11;q12) commonly found in the 8p11 myeloproliferative syndrome. The cDNA encodes a predicted protein of 1377 amino acids that is widely expressed and is highly homologous to the DX6673E/KIAA0385 and KIAA0425 proteins of uncertain functions. However, DX6673E has been implicated in a form of X-linked mental retardation. Alignment of these proteins revealed a novel zinc finger-related motif (MYM domain) that is repeated five times in each protein and closely resembles structures involved in mediating homo- and heterotypic protein-protein interactions. In patients with the t(8;13)(p11;q12) myeloproliferative syndrome , the *ZNF198* gene was fused to exon 9 of the FGF receptor-1 (*FGFR1*) that maps to 8p11. The authors hypothesize that this fusion leads to constitutive activation of the FGFR1 tyrosine kinase

signal transduction pathway that might contribute to malignant transformation (Popovici *et al.*, 1998; Reiter *et al.*, 1998; Smedley *et al.*, 1998; Xiao *et al.*, 1998). Recently, characterization of the genomic structure of *ZNF198* revealed the presence of 26 exons with the translation initiation codon located in exon 4 (Kulkarni *et al.*, 1999).

Tubulin is the structural subunit protein of microtubules that are essential to a variety of functions including cell motility, cell shape and polarity, transport and mitosis. It is composed of two 50 kD polypeptides designated α and β that each represents a family of isoforms generated by multiple tubulin genes and by posttranslational modifications. In mammals, there are at least five isotypes of α and six of β tubulins that are expressed in specific cell types or at discrete stages of development. The heterogeneity of tubulin proteins contributes to the diversity of microtubule function through either differential polymerization of the various tubulin subunits or by unique interactions with microtubule-associated proteins (Murphy, 1991; Luduena *et al.* , 1992; Raff, 1994; Mandelkow and Mandelkow, 1995). *TUBA2* is an α-tubulin gene isolated as part of positional cloning efforts to identify the gene(s) responsible for two forms of human nonsyndromic deafness (DFNA3 and DFNB1) that map to 13q11. The gene contains five exons and encodes a putative protein of 450 amino acids highly homologous to mouse testis α-tubulins. Northern blot analysis revealed a 1. 8 kb band in the human testis and a 1.1 kb band in the spleen, small intestine, colon, thymus, and lymphocytes. Southern blot analysis and direct sequencing excluded *TUBA2* as the gene responsible for DFNA3 and/or DFNB1 (Dodé *et al.*, 1998). As mentioned previously, these two diseases were found to be caused by mutations in connexin 26 (Kelsell *et al.*, 1997; Kelley *et al.*, 1998).

Cx26, Cx46, FGF-9, INF198 and *TUBA2* were considered as candidates for HED mainly because of their colocalization to the same candidate region. However, even though only one gene, *Cx26*, was shown to be involved in skin and hair development, the other four genes could be involved in this complex process whose underlying molecular events are poorly understood. The mechanisms by which the gap junction compartments in skin are established are not well characterized and other connexins including Cx46 could be required for this type of communication. FGF-9 could have a crucial regulatory role only at a specific time of morphogenesis of the epidermal appendages. Which isoform of α-tubulin is expressed in the hair/nail is presently unknown leaving the possibility that TUBA2 could be an important structural cellular component of these appendages. *ZNF198* is a novel gene with structural motifs most likely involved in mediating protein-protein interaction so this gene could contribute to a wide range of cellular functions. In this study, we describe the cloning of human *Cx46* coding sequence and the analysis of these five candidate genes for involvement in HED by polymerase chain reaction (PCR) and single strand conformation polymorphism (SSCP).

MATERIALS AND METHODS

Isolation of the human *CX46* coding sequence

Twelve primer pairs were designed from the rat connexin 46 (*Cx46*) cDNA sequence (Paul *et al.*, 1991; GeneBank Accession No. X57970) using PrimerSelect (DNASTAR) (data not shown). These primers were used to amplify from human genomic DNA with the following conditions: one cycle of denaturation at 94°C for 5 minutes followed by 35 cycles each consisting of 30 sec at 94°C, 30 sec at the corresponding

annealing temperature and 30 sec at 72°C followed by a last cycle of elongation at 72°C for 5 min. PCR reactions were carried out in a total volume of 17 µl containing: 200 ng of genomic DNA template, 125 ng of each primer, 200 µM each of dATP, dCTP, dTTP and dGTP, 0.5 U *Taq* DNA polymerase (Perkin Elmer) and 1.25 µl *Taq* buffer with 15 mM $MgCl_2$ (Perkin Elmer). Two primer pairs (5'-GCAAGGTGTGGCTGACCGTCCTG 3' and 5' TTGCAGCGCCCAGAAGCGAATGT 3'; 5' CGCGCTGCTGCGGACCTACG 3' and 5' TGACGCACAGGCCACAGCCAACAT 3') amplified the same size PCR fragments (181 bp and 214 bp respectively) in total human genomic DNA, chromosome 13-only cell hybrid (GM11689), hamster and rat DNA. These two PCR products were amplified in a 100 µl reaction, separated on 0.8% low melting point agarose, purified and used as probes against southern blots containing *Eco*RI or *Hin*dIII-digests of a set of 57 PACs that partially cover the HED candidate region. Southern blot analysis was done by standard methods (Sambrook *et al.*, 1989) except that gene- specific primers were used for labelling. Positive signals were obtained with ~ 4.7 Kb *Hin*dIII and >12 Kb *Eco*RI bands from the PAC 20G12.

To subclone the 4.7 Kb *Hin*dIII fragment, an arrayed plasmid library of the PAC 20G12 was constructed. The PAC clone was grown overnight in Terrific Broth (TB) supplemented with 12.5 µg/ml kanamycin and DNA was extracted using a modified alkaline-lysis technique (Birnboim *et al.*, 1979). One microgram of PAC DNA was digested with *Hin*dIII and blindly subcloned into pUC18 *Hin*dIII /BAP (Pharmacia Biotech). White colonies were arrayed into a 96-well tissue culture plate (Falcon 3072) containing TB with 60 µg/ml Ampicillin, incubated overnight at 37°C and then transferred to Hybond-N^+ nylon membrane (Amersham) using a 96-pin replicating tool.

The blot was hybridized against the Cx46 214 bp probe, positive clones were identified and DNA was prepared from overnight cultures in TB supplemented with 60 μg/ml Ampicillin by the alkaline-lysis method. Plasmid DNA was sequenced by the dideoxy chain termination method manually using a Sequenase Kit (USB) or with a Li-Cor DNA sequencer 4200. DNA sequence assembly was done using SeqMan (DNASTAR).

HED subjects

One to six unrelated HED affected individuals, each of different ethnic origin (French Canadian, Indian, Scottish-Irish, French, Spanish and South African), were available at different stages of this study. They presented with various degrees of nail dysplasia, hair defects and palmoplantar hyperkeratosis.

Genomic DNA and RNA isolation

Around 20 ml blood were obtained from individuals of French Canadian, French and Spanish origins: 10 ml were used to establish Epstein-Barr virus transformed lymphoblastoid cell lines and the other 10 ml were used for direct DNA extraction using standard procedures. RNA was isolated form the cell lines using Trizol (Life Technologies, Inc.) as suggested by the manufacturer. For individuals of Indian, Scottish-Irish and South African origin, only DNA was available. In addition, DNA and RNA from 2-5 normal controls were included in the SSCP analysis.

Reverse transcription (RT) and polymerase chain reaction (PCR)

Approximately 1 μg total RNA was reverse transcribed and PCR-amplified using the GeneAmp kit (Perkin Elmer) as recommended by the manufacturer, except that 1μl of 10mM dATP was substituted with 12.5 μCi [^{35}S]dATP in the Master mix for reverse transcription. Genomic DNA was PCR-amplified in a total volume of 17 μl containing: 100-200 ng of genomic DNA; 200 μM of dCTP, dGTP and dTTP; 25 μM dATP, 125 ng of each primer; 2.5 μCi [^{35}S]dATP, 0.5 U Taq DNA polymerase (Perkin Elmer) and 1. 25 μl Taq buffer with 15 mM $MgCl_2$ (Perkin Elmer). PCR conditions for amplification of reverse transcribed RNA and genomic DNA consisted of one cycle of denaturation for 5 minutes at 94°C followed by 35 cycles with each cycle consisting of 30 sec at 94°C, 30 sec at the corresponding annealing temperature (A. T. °C, see table 3. 1) and 30 sec at 72°C followed by one cycle of elongation for 5 min at 72°C.

PCR primers were designed using Primer Select (DNASTAR) to amplify overlapping sequences of the open reading frame (ORF) of each of the five candidate genes: *Cx26* (Genebank Accession No. M86849), *Cx46* (cloned in our laboratory), *TUBA2* (Genebank Accession No. AF005392), *FGF-9* (Genebank Accession No. D14838) and *ZNF198* (Genebank Accession No. AJ224901). In addition, PCR primers that flank the intron/exon boundaries were developed for the gene *TUBA2* whose genomic structure is well characterized (Dodé *et al.*, 1998). For *ZNF198*, primers numbered 44 and less were generously provided by Dr. N. C. Cross from the Imperial College School of Medicine in U.K. PCR products larger than 300 bp were digested with the appropriate enzyme prior to electrophoresis. Table 3.1. lists the primer sequences,

Table 3. 1. PCR primers for SSCP analysis of Cx26, Cx46, FGF-9, TUBA2 and ZNF198

Gene	Primer pair	Sequence (5'->3')	A. T. °C	Size (bp)/enzyme
Cx26	GJB2.1	AGAGCAAACCGCCCAGAGTAGAAG	58	247
		GGATGTGGGAGATGGGGAAGTAGT		
	GJB2.2	AGATGAGCAGGCCGACTTTGTCT	58	197
		TCTCCCCCTTGATGAACTTCCTCT		
	GJB2.3	CACGTGGCCTACCGGAGACAT	56	170
		AAGGCGGCTTCGAAGATGACC		
	GJB2.4	AAGGCTCCCTGTGGTGGACCTAC	58	169
		GGCCGGGACACAAAGCAGTC		
	GJB2.5	CAACGCCTGGCCTTGTCC	56	234
		ACGGGTTGCCTCATCCCTCTC		
Cx46	MRcons1	TGCGGACCCGGCACTCAGC	60	194
		GCGGCCCCCAGCACCAAGAT		
	CX46cons-1	CACTCCACGGTCATCGGCAAGGTT	58	246
		GTGCAGCACGTGGCCCAGGTAGA		
	CX46cons-2	CGCTGCAGATCATCTTCGTGTCCA	58	250
		CCCACCTCGAACAGCGTCTTGAAG		
	Cx46-MF2-1	CGGGGCGCTGCTGCGAACCT	55	206
		CCACCGCCAGCATGAAGAT		
	CX46cons-3	CCTGCCCCAACACGGTGGACTG	64	268
		TGTGCGCATAGTAGGGTGGGAACC		
	CX46cons-4	CCCTGCCCCCCAGCTCCC	67	248
		GCTCGGCCGCCTGGTTGG		
	CX46cons-5	AACGGCCACCACCACCTGCTGAT	63	202
		CCAGACTGCTGCCGCTCCCATCCA		
	Cx46-MF1-1	TCGCGCACGAGGCTGAG	62	230
		CCGGGCACTAGATGGCCAAGTCCT		
	Cx46-MF1-2	CGCCCTTGCCCCTCGGAGAC	55	190
		CTGGGACTTTCAGGTTCTATCTGC		
	Cx46-MF1-3	TGCCTCCAGACAGCTGTGTAGTGA	55	227
		GTCCCCGCCACCCCCAAACTC		
FGF-9	FGF9-1	CCTGGGTTGACACCATCATTATTG	58	188
		CCCCCTGCTTCGGACTGACC		
	FGF9-2	CGGTGGACAGCCCGGTTTTGTTA	58	198
		ATCGGCTGTGGTCTTTCCTGGTTC		

Table 3. 1. Continued.

FGF-9	FGF9-3	CAGGGGACCCGCAGTCACG ACAGCTCCCCCTTCTCATTCATCC	64	236
	FGF9-4	AGGCGTGGACAGTGGACTCTACCT TGAATTTCTGGTGCCGTTTAGTCC	65	230
	FGF9-5	CGTGGACACTGGAAGGCGATACTA GAACCCACCGCGTGAAACCTTTAT	55	230
TUBA2	TUBA2 EX1-2	GGTGGGGACGACTCCTTCAA ATGGGGCCTTACCGACGAT	58	165
	TUBA2 EX2	GGTGAGGATAGGCCTTAAAGATTC CCTGTGCAGGAAAATGGTCAC	56	238
	TUBA2 EX3-1	CAGAAGTGGCCCTGGCTTGGTG TTCTTGCCGTAATCCACTGAGAGC	58	194
	TUBA2 EX3-2	AGTTTTGGGGGTGGCACTG GTCCAGGTTGCGCCGACATA	57	237
	TUBA2 EX3-3	CAACGCTGGAACATTCTGACTGTG GGATGCGGGGGTACGGCACTA	58	219
	TUBA2 EX3-4	CACGGCCTCCCTGCGATTTG TACTTGCCGTGGCGAGGGTCAC	60	222
	TUBA2 EX3-5	CTTCGAGCCAGCCAATCAGATGGT GGGCCTTCAGGGGCAGCATTAC	58	250
	TUBA2 EX4-1	AAAGCAGCCACCATTTCCAGGTTT ACAAAGGCCCGCTTGGCATACA	60	225
	TUBA2 EX4-2	CTGGGCTCGCCTGGACCATAAGTT GCTGGGGGTGGCAGTGGAGTG	60	236
ZNF198	DET 50/51	ACACCCCCATTGTTAGATGAGGT TCTTCCACTGACGAGTTCTGAAAC	53	235
	DET 52/53	CATGGCAACTAGTCTCACGAATGT AATTTTTCTCTTGCCCTTGATTTG	53	335/TaqI
	DET 15/14	TTGGCATCTCAGAAGGGAAGTG TCATTCACAGAGCTTACACCTGT	52	281
	DET 54/55	AGGAATGGGTAATAGTGGTATCAC TGAAGTTAAAGATGGGGTGTATGT	52	250
	DET 56/8	TTGGGAGATGTCTCTAACGGACTG TCCACCCCTGGTTGTTTCTGAT	54	233
	DET 57/58	CCTGATTCTTGGATCTCCCAGTCA GGCAGGTGGTAGAACAAAAGAGGT	57	235

Table 3. 1. Continued.

ZNF198	DET 59/60	AGGGCCAGACAGCTTATCAACGAA TGGCGAATCTCAGTTAGTTTACCA	53	289
	DET 61/12	TCTCTATGAAGACAAACAGAATCC GTTGACCATCAATCACCAGAAC	50	248
	DET 62/44	GTACTTGCCCAGTAAAGGTGC AGAGTCCTGCATGTGATCTCGA	52	151
	DET 63/29	TACAAACAGGTAGGTAGCCATCCA AGGACCTATACACTGAGTCAT	50	159
	DET 22/64	TTGTACTGGTTGCCGAACACAG GTACAGTTTTCCATCAGGCATTGT	52	211
	DET 65/4	AGCGAAACTCTTTACCTCAATACC CTGAATATCACTTGGCGCTACA	52	146
	DET 3/66	GTGGCTAAATTTCAGGCTCT ATTCGCAATATGTAACTATGC	50	211
	DET 67/68	GTTCTGCAGCAAAACTTGTTCAG TTAGTTGCTCCCTTCTTACATAGC	52	237
	DET 69/70	ATTTTGCCAGACGTTTAGGATTGA GTCATGTTGGGCTCATTTTGTTGA	54	283
	DET 71/40	GTACTACAAGGCTGCAAGGTGTG CAGGGATAGGCACTGGAACATA	54	383/TaqI
	DET 72/73	TCACATGCAGACCAAATCTTGTCA TGTTTTCCCCTCGTCTTCACTCAT	56	298
	DET 74/75	AAAAAGCAAGGTTTCTTCAGATGC TTCGCCAAAAACAGGTGGTAATAA	52	283
	DET 41/76	AGAGCTGCTGAGGAGCTTGATA TTGCCTAGTTTTGACCCAGTGTTT	54	228
	DET 77/78	TTCAGAATGCAGCTTTCCTTTCAA CATTTGGCCGTCGGATCTCATTG	54	212
	DET 79/18	AACTATGGGTTAGCTCATTTTGT TATGCTTGGTTGCCAGCTTCGCA	50	204
	DET 80/24	AACATTTGAGCAAGAATTGAATAA CACTTGGTATCGAAGACACGCT	52	286
	DET 81/82	TTTGGCCTGAAAACAGTGG AGACAAGTAGCATTCAAACATT	50	258
	DET 83/38	CAGCCAATCCTTCCAGATGTCCTG TTTATCCCGATTGCTTCTGCAACGT	54	279

annealing temperature (A. T.°C), PCR product size and restriction enzyme (when used) for each of the five genes screened.

SSCP analysis

The SSCP technique detects sequence changes as shifts in electrophoretic mobility. Under nondenaturing conditions, the single-stranded DNA has a three-dimensional conformation that is dependent on its sequence. Any change in this sequence, including single-base substitutions, will lead to an alteration in the conformation of the strand and hence its mobility. The mutant strand will migrate differently on a nondenaturing polyacrylamide gel and will be detected as a variant band compared to controls (Orita *et al.*, 1989).

The PCR products were mixed with 5 μl formamide-containing loading buffer, denatured for 5 minutes at 90°C and loaded on 6-9% nondenaturing poyacrylamide gel with or without 5% glycerol. Gels were run at 4°C in 1×TBE at 21 Watts for 4-6hrs, dried and exposed to BioMax MR X-ray films (Kodak) for 16-24 hrs at room T°C.

RESULTS

Sequence of the ORF of human *Cx46*

We have used the rat *Cx46* cDNA sequence to clone homologous sequences in human DNA. As described in Materials and Methods, using consensus primers to rat *Cx46*, a 214 bp PCR fragment was amplified from chromosome-13 only hybrid cell line and hybridized with a 4.7 Kb Hind III fragment isolated from PAC 20G12 that maps to the region. Consequently, this fragment was subcloned and sequenced. Alignment and

comparison of the obtained sequence with the mouse cDNA *Cx46* allowed us to determine the initiation and stop codons of the ORF. The ORF spans around 1.3 Kb and is most likely contained all in one exon, similar to all other characterized members of the connexin family (Bruzzone *et al.*, 1996). At the same time, Mackay *et al.* (1997) have mapped a form of autosomal dominant cataract to the same region as HED and were trying to isolate the human *Cx46* for mutation screening in the cataract patients. Using consensus primers to rodent cDNA sequences and human *Cx50*, they identified the 5' coding sequence for *Cx46* but were unable to amplify the remaining 3' coding sequence. Pooling of sequencing data from both groups detected a single ORF coding for a predictive protein of 435 amino acids with a calculated molecular mass of 47,435 Da. The human *Cx46* sequence shared ~88% homology and ~70% identity with that of the mouse, particularly in the transmembrane and extracellular domains of the protein (MacKay *et al.*, 1999).

SSCP analysis

The ORF of each of the genes, *Cx26*, *Cx46*, *FGF-9*, *ZNF198*, and *TUBA2*, was screened for the presence of mutations in HED patients by PCR-SSCP analysis. For *TUBA2*, the analysis included the exon-intron junctions as well and did not include the codons specifying the first 50 amino acids of the protein because the available genomic sequence lacks the first exon along with the initiator methionine codon and our PCR attempts failed to amplify the first 147 bp of the ORF. SSCP analysis was carried out on amplified genomic DNA with *Cx26* and *Cx46*, whose ORF each resides in one exon, and *TUBA2*, and mainly on RT-PCR products with *FGF-9* and *ZNF198*, except for FGF9-1,

FGF9-2, FGF9-4, DET 52/53 and DET 54/55. No mutation specific to HED was detected in any of the five genes tested suggesting that these genes are most likely not defective in HED. We encountered few non-pathogenic polymorphisms, four of which, obtained with the primers MF1-3 (*Cx46*), DET 52/53 (*ZNF198*), TUBA2 EX1-2 and TUBA2 EX3-5, could be informative and hence useful as genetic markers in the region. In figure 3.1., part a illustrates an example of SSCP results with no migrational difference between affected and controls DNA indicating the absence of basepair changes in the sequence tested and part B shows an example of normal SSCP polymorphism where a variant band is seen in some affected and control individuals.

DISCUSSION

Base substitutions, small deletions and insertions, are among the most common mutations occurring in the coding sequence of the gene, constituting the majority of the pathogenic changes so far identified. Many methods have been developed to detect these types of mutations (reviewed in Grompe, 1993), the most commonly used of which is SSCP. This method is relatively simple and highly sensitive with the ability to detect 70-95% of mutations in fragments no larger than 200 bp. However, this sensitivity decreases with the size of the PCR product and reaches less than 50% when fragments analyzed are >400 bp. In the present study, the size of PCR products analyzed ranged from 146 to 383 bp with only 2 products larger than 300 bp. These two products were restriction digested before being analyzed to overcome the size limitations of the technique, avoiding the loss of informativeness associated with band size. In addition, the likelihood of finding a mutation was increased by testing more than one individual predicted to carry a different

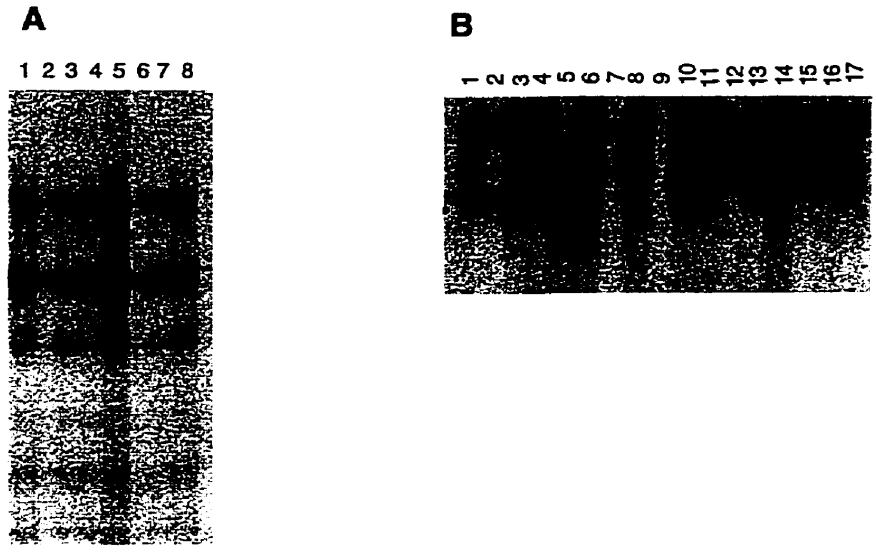

Figure 3. 1. Examples of PCR-SSCP results. A) lanes 1-4: controls; lanes 5-8: affected. All bands have the same migrational pattern indicating the absence of basepair changes in the sequence tested. B) lanes 1-5: controls; lanes 6-17: affected. A variant band is seen in both controls and affected individuals indicating the presence of a normal SSCP polymorphism. Individuals in lanes 1, 11, 13-15 and 17 are heterozygous at this polymorphic locus. Individuals in lanes 2 and 3 are homozygous for one allele and those in lanes 4, 6, 8, 10, 12 and 16 are homozygous for the other allele. Individual in lane 5 most likely carries a basepair alteration at another site in the sequence amplified. In lanes 7 and 9, PCR did not work.

mutation of the HED gene. In fact, a panel of 3-6 unrelated affected individuals, each of different ethnic origin and carrying a distinct haplotype, was available for mutation screening of the candidate genes (except for *Cx26* where only French Canadian individuals were tested).

Mutation screening of these candidate genes was carried out in parallel to construction of physical and genetic maps of the HED candidate region, collection of additional HED families and isolation of new polymorphic markers to further refine the localization of the HED gene by linkage and linkage disequilibrium analyses. Our efforts have led to narrowing down the HED candidate region from >15 cM to 2.4 cM flanked by D13S1828 proximally and D13S1830 distally (Kibar *et al.*, submitted). These new boundaries exclude *FGF-9* from the candidate region but still include *Cx26*, *Cx46*, *ZNF198* and *TUBA2*.

After we tested and excluded the ORF of *Cx26* and *Cx46* by PCR-SSCP for the presence of pathogenic mutations in HED, mutations in these two genes were found in hereditary non-syndromic sensorineural deafness (Kelsell *et al.*, 1997) and autosomal dominant "zonular pulverulent" cataract (Mackay *et al.*, 1999) respectively. Kelsell et al. (1997), in the same study, reported the exclusion of *Cx26* in HED by finding no sequence variants of the coding exon in a family affected with HED thereby supporting our SSCP findings.

Our SSCP analysis of *ZNF198* and *TUBA2* in HED suggests that most likely these two genes are not responsible for HED. However, we have only screened the ORF of *ZNF198* and around 90% the ORF and splice sites of *TUBA2*. Pathogenic mutations can also occur in the intronic sequences, the 5′ and 3′ untranslated regions and regulatory

sequences usually located in the promoter region (Strachan and Read, 1997) and so there remains a small chance that one of them is involved in HED. The candidate region of 2.4 cM is still large for positional cloning strategies and further efforts are being conducted to minimize it. Detailed analysis of these two genes for involvement in HED will be carried out if they will still map to the candidate region once additional genetic studies further reduce the size of the region.

ACKNOWLEDGMENTS

We would like to thank the families that participated in this study. We thank Dr. Nicholas Cross for sending us the cDNA sequence and PCR primers for ZNF198. We are grateful to the CGAT Genome Resource Facility at the Hospital for Sick Children in Toronto for providing PAC and YAC clones. This work was supported by grants from the Canadian Dermatology Foundation and the National Foundation for Ectodermal Dysplasias. G.A.R is supported by the Medical Research Council of Canada.

REFERENCES

Birnboim, H. C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acid Res.* **7:** 1513-1523.

Bruzzone, R., White, T. W., and Paul, D. L. (1996). Connections with connexins: the molecular basis of direct intercellular signaling. *Eur. J. Biochem.* **238:** 1-27.

Butterweck, A., Elfgang, C., Willecke, K., and Traub, O. (1994). Differential expression of the gap junction proteins connexin45, -43, -40, -31, and -26 in mouse skin. *Eur. J. Cell Biol.* **65:** 152-163.

Dodé, C., Weil, D., Levilliers, J., Crozet, F., Chaïb, H., Levi-Acobas, F., Guilford, P., and Petit, C. (1998). Sequence characterization of a newly identified human alpha-tubulin gene (TUBA2). *Genomics* **47:** 125-130.

Du Cros, D. L. (1993). Fibroblast growth factor and epidermal growth factor in hair development. *J. Invest. Dermatol.* **101(1 Suppl):** 106S-113S.

Galzie, Z., Kinsella, A. R., and Smith, J. A. (1997). Fibroblast growth factors and their receptors. *Biochem. Cell Biol.* **75:** 669-685.

Goliger, J. A., and Paul, D. L. (1994). Expression of gap junction proteins Cx26, Cx31.1, Cx37, and Cx43 in developing and mature rat epidermis. *Dev. Dyn.* **200:** 1-13.

Gong, X., Li, E., Klier, G., Huang, Q., Wu. Y., Lei, H., Kumar, N. M., Horwitz, J., and Gilula, N. B. (1997). Disruption of alpha3 connexin gene leads to proteolysis and cataractogenesis in mice. *Cell* **91:** 833-843.

Goodenough, D. A., Goliger, J. A., and Paul, D. L. (1996). Connexins, connexons, and intercellular communication. *Annu. Rev. Biochem.* **65:** 475-502.

Grompe, M. (1993). The rapid detection of unknown mutations in nucleic acids. *Nature Genet.* **5:** 111-117.

Guilford, P., Dode, C., Crozet, F., Blanchard, S., Chaib, H., Levilliers, J., Levi-Acobas, F., Weil, D., Weissenbach, J., Cohen, D., Le Paslier, D., Kaplan, J-C., and Petit, C. (1995). A YAC contig and an EST map in the pericentromeric region of chromosome 13 surrounding the loci for neurosensory nonsydromic deafness (DFNB1 and DFNA3) and limb-girdle muscular dystrophy type 2C (LGMD2C). *Genomics* **29:** 163-169.

Hébert, J. M., Rosenquist, T., Götz, J., and Martin, G. R. (1994). FGF5 as a regulator of the hair growth cycle: evidence from targeted and spontaneous mutations. *Cell* **78:** 1017-1025.

Kamibayashi, Y., Oyamada, M., Oyamada, Y., and Mori, M. (1993). Expression of gap junction proteins connexin 26 and 43 is modulated during differentiation of keratinocytes in newborn mouse epidermis. *J Invest Dermatol.* **101:** 773-778.

Kelley, P. M., Harris, D. J., Comer, B. C., Askew, J. W., Fowler, T., Smith, S. D., and Kimberling, W. J. (1998). Novel mutations in the connexin 26 gene (GJB2) that cause autosomal recessive (DFNB1) hearing loss. *Am. J. Hum. Genet.* **62:** 792-799.

Kelsell, D. P., Dunlop, J., Stevens, H. P., Lench, N. J., Liang, J. N., Parry, G., Mueller, R. F., and Leigh, I. M. (1997). Connexin 26 mutations in hereditary non-syndromic sensorineural deafness. *Nature* **387:** 80-83.

Kettunen, P., and Thesleff, I. (1998). Expression and function of FGFs-4, -8, and -9 suggest functional redundancy and repetitive use as epithelial signals during tooth morphogenesis. *Dev. Dyn.* **211:** 256-268.

Kibar, Z., Der Kaloustian, V. M., Brais, B., Hani, V., Fraser, F. C., and Rouleau, G. A. (1996). The gene responsible for Clouston hidrotic ectodermal dysplasia maps to the pericentromeric region of chromosome 13q. *Hum. Mol. Genet.* **5:** 543-547.

Kibar, Z., Dubé, M. –P., Powell, J., McCuaïg, C., Hayflick, S. J., Zonana, J., Hovnanian, A., Radhakrishna, U., Antonarakis, S. E., Benohanian, A., Sheeran, A. D., Stephan, M. L., Gosselin, R., Kelsell, D. P., Christianson, A. L., Fraser, F. C., Der Kaloustian, V. M., and Rouleau, G. A. Clouston hidrotic ectodermal dysplasia (HED): genetic

homogeneity, presence of a founder effect in the French Canadian population and fine genetic mapping. Submitted.

Kulkarni, S., Reiter, A., Smedley, D., Goldman, J. M., and Cross, N. C. P. (1999). The Genomic Structure of ZNF198 and Location of Breakpoints in the t(8;13) Myeloproliferative Syndrome. *Genomics* **55:** 118-121.

Kumar, N. M., and Gilula, N. B. (1992). Molecular biology and genetics of gap junction channels. *Semin. Cell Biol.* **3:** 3-16.

Lee, S. W., Tomasetto, C., Paul, D., Keyomarsi, K., and Sager, R. (1992). Transcriptional downregulation of gap-junction proteins blocks junctional communication in human mammary tumor cell lines. *J. Cell Biol.* **118:** 1213-1221.

Luduena, R. F., Banerjee, A., and Khan, I. A. (1992). Tubulin structure and biochemistry. *Curr. Opin. Cell Biol.* **4:** 53-57.

Mackay, D., Ionides, A., Berry, V., Moore, A., Bhattacharya, S., and Shiels, A. (1997). A new locus for dominant "zonular pulverulent" cataract, on chromosome 13. *Am. J. Hum. Genet.* **60:** 1474-1478.

Mackay, D., Ionides, A., Kibar, Z., Rouleau, G., Berry, V., Moore, A., Shiels, A., and Bhattacharya, S. (1999). Connexin 46 mutations in autosomal dominant congenital cataract. *Am. J. Hum. Genet* **64:** 1357-1364.

Mandelkow, E., and Mandelkow, E. M. (1995). Microtubules and microtubule-associated proteins. *Curr. Opin. Cell Biol.* **7:** 72-81.

Mattei, M. G., Penault-Llorca, F., Coulier, F., and Birnbaum, D. (1995). The human FGF9 gene maps to chromosomal region 13q11-q12. *Genomics* **29:** 811-812.

Mignon, C., Fromaget, C., Mattei, M. G., Gros, D., Yamasaki, H., and Mesnil, M. (1996). Assignment of connexin 26 (GJB2) and 46 (GJA3) genes to human chromosome 13q11-->q12 and mouse chromosome 14D1-E1 by in situ hybridization. *Cytogenet. Cell Genet.* **72:** 185-186.

Mitsui, S., Ohuchi, A., Hotta, M., Tsuboi, R., and Ogawa, H. (1997). Genes for a range of growth factors and cyclin-dependent kinase inhibitors are expressed by isolated human hair follicles. *Br. J. Dermatol.* **137:** 693-698.

Miyamoto, M., Naruo, K. –I., Seko, C., Matsumoto, S., Kondo, T., and Kurokawa, T. (1993). Molecular cloning of a novel cytokine cDNA encoding the ninth member of the fibroblast growth factor family, which has a unique secretion property. *Mol. Cell Biol.* **13:** 4251-4259.

Murphy, D. B. (1991). Functions of tubulin isoforms. *Curr. Opin. Cell Biol.* **3:** 43-51.

Nicholson, S. M., and Bruzzone, R. (1997). Gap junctions: getting the message through. *Curr. Biol.* **7:** R340-R344.

Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. (1989). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5:** 874-879.

Paul, D. L.(1995). New functions for gap junctions. *Curr. Opin. Cell Biol.* **7:** 665-672.

Paul, D. L., Ebihara, L., Takemoto, L. J., Swenson, K. I., and Goodenough, D. A. (1991). Connexin46, a novel lens gap junction protein, induces voltage-gated currents in nonjunctional plasma membrane of Xenopus oocytes. *J. Cell Biol.* **115:** 1077-1089.

Peus, D., and Pittelkow, M. R. (1996). Growth factors in hair organ development and the hair growth cycle. *Dermatol. Clin.* **14:** 559-572.

Popovici, C., Adélaïde, J., Ollendorff, V., Chaffanet, M., Guasch, G., Jacrot, M., Leroux, D., Birnbaum, D., and Pébusque, M. -J. (1998). Fibroblast growth factor receptor 1 is fused to FIM in stem-cell myeloproliferative disorder with t(8;13). *Proc. Natl. Acad. Sci. USA.* **95:** 5712-5717.

Raff, E. C. (1994). The role of multiple tubulin isoforms in cellular microtubule function. *Microtubules*: 85-109.

Reiter, A., Sohal, J., Kulkarni, S., Chase, A., Macdonald, D. H., Aguiar, R. C., Goncalves, C., Hernandez, J. M., Jennings, B. A., Goldman, J. M., and Cross, N. C. (1998). Consistent fusion of ZNF198 to the fibroblast growth factor receptor-1 in the t(8;13)(p11;q12) myeloproliferative syndrome. *Blood* **92:** 1735-1742.

Risek, B., Klier, F. G., and Gilula, N. B. (1992). Multiple gap junction genes are utilized during rat skin and hair development. *Development* **116:** 639-651.

Rosenquist, T. A., and Martin, G. R. (1996). Fibroblast growth factor signalling in the hair growth cycle: expression of the fibroblast growth factor receptor and ligand genes in the murine hair follicle. *Dev. Dyn.* **205:** 379-386.

Salomon, D., Masgrau, E., Vischer, S., Ullrich, S., Dupont, E., Sappino, P., Saurat, J. H., and Meda, P. (1994). Topography of mammalian connexins in human skin. *J. Invest. Dermatol.* **103:** 240-247.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) In *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor laboratory Press. **Vol. 2**, pp. 9.31-10.16.

Smedley, D., Hamoudi, R., Clark, J., Warren, W., Abdul-Rauf, M., Somers, G., Venter, D., Fagan, K., Cooper, C., and Shipley, J. (1998). The t(8;13)(p11;q11-12) rearrangement associated with an atypical myeloproliferative disorder fuses the

fibroblast growth factor receptor 1 gene to a novel gene RAMP. *Hum. Mol. Genet.* **7:** 637-642.

Strachan, T., and Read, A. P. (1997). Mutation and instability of human DNA. In *Human Molecular Genetics*. BIOS Scientific Publishers limited, Wiley-Liss, pp. 241-273.

White, T. W., Bruzzone, R., and Paul, D. L. (1996). Connections with connexins: the molecular basis of direct intercellular signaling. *Eur. J. Biochem.* **238:** 1-27.

Wilkie, A. O. M., Morriss-Kay, G. M., Jones, E. Y., and Heath, J. K. (1995). Functions of fibroblast growth factors and their receptors. *Curr. Biol.* **5:** 500-7.

Xiao, S., Nalabolu, S. R., Aster, J. C., Ma, J., Abruzzo, L., Jaffe, E. S., Stone, R., Weissman, S. M., Hudson, T. J., and Fletcher, J. A. (1998). FGFR1 is fused with a novel zinc-finger gene, ZNF198, in the t(8;13) leukaemia/lymphoma syndrome. *Nat. Genet.* **18:** 84-87.

CHAPTER 4

A radiation hybrid map of 48 loci including the Clouston hidrotic ectodermal dysplasia (HED) locus in the pericentromeric region of chromosome 13q

This radiation hybrid map was constructed to provide order and distance estimates for loci across the HED candidate region and to serve as a basis for isolation of candidate genes and new polymorphic markers from the region

The presented work was published in a shortened version as requested by the editors of *Genomics* as:

Kibar, Z., Lafrenière, R. G., Chakravarti, A., Wang, J. -C., Chevrette, M., Der Kaloustian, V.M., and Rouleau, G. A. A radiation hybrid map of 48 loci including the Clouston hidrotic ectodermal dysplasia (HED) locus in the pericentromeric region of chromosome 13q (1999). *Genomics* **56:**127-130.

I chose to include the more detailed version of the paper to make the chapter easier to comprehend.

Contribution by co-authors:

Kibar: PCR typing of the radiation hybrid clones; analysis of radiation hybrid data; isolation and characterization of new polymorphic markers

Lafrenière: helped isolate the new polymorphic markers

Chakravarti: provided the radiation hybrid clones

Wang and Chevrette: Fiber-FISH and FISH on interphase nuclei

Der Kaloustian: helpful comments

Rouleau: lab. supervisor

ABSTRACT

Hidrotic ectodermal dysplasia Clouston type (HED) is an autosomal dominant skin disorder which maps to the pericentromeric region of chromosome 13q. To facilitate the identification of the HED gene, we used a chromosome 13-specific radiation hybrid panel to map 54 loci in the HED candidate region; the marker retention data were analyzed using RHMAP version 3. The 54 markers have an average retention frequency of 31.6 % with decreasing retention as a function of distance from the centromere. Two-point analysis identified three linkage groups with a threshold lod score of 4.00: one linkage group consisted of 49 loci including the centromeric marker D13Z1 and the telomeric flanking marker for the HED candidate region D13S143. Assuming a centromeric retention model, multi-point maximum likelihood analysis, of these 49 loci except D13Z1, provided a 1000:1 framework map ordering 29 loci with 21 unique map positions and ~2000 times more likely than the next order. Loci that could not be ordered with this level of support were positioned within a range of adjacent intervals. This map spans 347 cR_{9000} and has an average resolution of 17.3 cR_{9000}, and includes 3 genes (TUBA2, GJβ2 and FGF-9), 18 ESTs, 19 polymorphic loci and 8 single copy DNA segments. Comparison of our RH map to a YAC contig showed an inconsistency in order involving a reversed interval of 6 loci. Fiber-FISH and FISH on interphase nuclei analyses with PACs isolated from this region supported our order. We also describe the isolation of eight new chromosome 13q polymorphic $(CA)_n$ markers that have an average PIC value of 0.67. These data and mapping reagents will facilitate the isolation of disease genes from this region.

INTRODUCTION

Clouston hidrotic ectodermal dysplasia (HED) is an autosomal dominant genodermatosis characterized by nail dystrophy, defects of the hair and hyperkeratosis of the palms and soles (Williams and Fraser, 1967; Der Kaloustian and Kurban, 1979). HED has been described in families of different ethnic origins but is most common in the French Canadian population (Clouston, 1929; Escobar *et al.*, 1983). Using linkage analysis in eight French Canadian families, we have mapped the gene responsible for HED to the pericentromeric region of chromosome 13q. Haplotype and multi-point linkage analyses have allowed us to refine the location of HED to a 12.7 cM region flanked distally by the marker D13S143 (Kibar *et al.*, 1996). Further refinement of the HED candidate region by genetic analysis requires the isolation of additional polymorphic markers and the construction of high resolution physical and genetic maps that can provide order and distance estimates for markers in the region.

Radiation hybrid (RH) mapping is a somatic cell genetic method proven to be valuable in constructing long-range physical maps of human chromosomes. In this method, a hybrid cell line containing a single human chromosome is lethally irradiated with X-rays breaking the human chromosome into smaller fragments. Subsequently, radiated cells are rescued by fusion to another recipient rodent cell line. Each of the resulting hybrid clones will have a unique set of human chromosome fragments and can be scored for the presence or absence of human chromosome-specific markers. The closer two loci are on a chromosome, the less likely they will be separated by an X-ray induced breakage. Based on this principle, analysis of the retention patterns of various loci in a set of RH clones can generate a physical map of the region in interest (Goss and Harris,

1975; Cox *et al.*, 1990). Various non-parametric and model-based analytical approaches have been proposed for map construction based on RH data (summarized in Boehnke, 1992; Lunetta and Boehnke, 1994). A commonly used method for multipoint analysis of RH data is the maximum likelihood method that can generate relative likelihoods of different marker orders and distance estimates between adjacent markers. This method is based on the assumption of independent fragment retention and random X-ray breakage along the chromosome (Boehnke *et al.*, 1991, Lunetta *et al.*, 1996). A major advantage of RH mapping versus genetic mapping is that it can map STSs derived from any type of DNA sequence, polymorphic or monomorphic, and therefore can integrate data from available physical and genetic maps. Furthermore, the level of resolution achieved by RH mapping is intermediate between that of genetic maps on one hand and other physical maps such as YAC contigs (Cox *et al.*, 1990).

In this study, we have used a set of 94 chromosome 13q RH clones to map 54 pericentromeric markers in the HED candidate region. RH data analysis allowed us to construct a radiation hybrid map of 48 loci including 3 genes, 18 ESTs, 19 polymorphic loci and 8 single copy DNA segments. In addition, we have isolated eight new polymorphic $(CA)_n$ repeat markers in the region that will aid in the cloning of the HED gene and other chromosome 13q pericentromeric disease genes.

MATERIALS AND METHODS

Radiation hybrid clones

A panel of 94 chromosome 13 radiation hybrid clones, as previously described by Shaw *et al.* (1995), was used. This panel was constructed by irradiating human-hamster

hybrid GM10898 cells, containing chromosome 13 as the only human material, at 9000 rads and fusing them with the HPRT deficient hamster cell line 380-6.

Chromosome 13q markers

A total of 54 chromosome 13q pericentromeric markers were used in this study, including eight new polymorphic markers. The other 46 loci were derived from the following maps or sites: the Genome Database (GDB) at http://gdbwww.gdb.org/, the Human Gene Map at http://www3.ncbi.nlm.nih.gov/science96/ (Schuler *et al.*, 1996), the 1996 Genethon comprehensive genetic map (Dib *et al.*, 1996), the MIT YAC contig WC13.0 at http://www-genome.wi.mit.edu/cgi-bin/contig/phys-map, a report of the third international workshop on human chromosome 13 mapping 1995 at http://genome1.ccc.columbia.edu/~genome/ (Warburton *et al.*, 1996), a somatic cell hybrid map of human chromosome 13 established by Still *et al.* (1996), a chromosome 13q radiation hybrid map developed by Shaw *et al.* (1995) and a YAC contig and an EST map constructed by Guilford *et al.* (1995) in the region surrounding the loci for neurosensory nonsyndromic deafness and limb-girdle muscular dystrophy type 2C.

PCR genotyping of the radiation hybrids

We scored RH clones for the presence or absence of the 54 markers by PCR. PCR conditions for most of the markers consisted of the following: one cycle of denaturation at 94°C for 5 minutes, followed by 35–38 cycles each consisting of 30 sec at 94°C, 30 sec. at the corresponding annealing temperature (A.T.) and 30 sec. at 72°C followed by a last cycle of elongation at 72°C for 5 min. The STS primers used for amplification and

the annealing temperatures were derived from the corresponding reference. We only used a different set of conditions with the markers D13S141, D13S143 (Pertrukhin *et al.*, 1993), D13S250 and D13S633 (GDB) as described. Table 4.1. lists the primer sequences, annealing temperature and the PCR product size range for each of the eight newly isolated $(CA)_n$ markers. Each PCR reaction was done in a 25 µl reaction containing the following: 100 ng of RH DNA, 250 ng of each primer, 280 µM dNTP, 1.5 U Taq DNA polymerase (Perkin Elmer) and 2.6 µl Taq buffer with 15 mM $MgCl_2$ (Perkin Elmer). The PCR products were separated on 2% agarose gels and visualized by ethidium bromide staining. The markers D13S1088E, D13S1236, D13S183E and D13S1316 were amplified with [^{35}S]dATP as described previously (Kibar *et al.*, 1996) for better resolution of human PCR products from the hamster background. Each marker was typed at least twice in the panel to reduce the probability of false positive and false negative scoring. All experiments included a positive control, the monochromosomal 13 hybrid GM10898, and a negative control, 380-6. The marker D13S1835 was scored only in 50 RH clones due to limited DNA availability.

Analysis of radiation hybrid data

The PCR radiation hybrid data were analyzed using the statistical package for multi-point radiation hybrid mapping RHMAP version 3 developed by Boehnke et al (Boehnke *et al.*, 1991). This method assumes that X ray breakage along the chromosome is random and that the resulting fragments are independently retained in a RH clone. We first used RH2PT for two-point analysis to estimate locus retention probabilities and to identify linkage groups and markers with identical retention patterns. Markers with

Table 4. 1. New Polymorphic $(CA)_n$markers on chromosome 13q

Marker (PAC)	Primer sequence (5'-3'; forward/reverse primer)	A.T. (°C)	Size range (bp)	Number of alleles	PIC
D13S1826 (72H5)[a]	TCATCTGAGAGACAGAGAGAG GCCAACATCCCTTCACACAAA	57	110-132	12	0.81
D13S1827 (207O19)	CCAGCCCAGGTGATAGACTGA TCTGGCCATTTGGTAGGTGTTA	60	181-195	8	0.43
D13S1828 (50F20)	TTGGGCGCTCAATCTAGTTACTT CACTGTGCCTGGTCCGATTTA	55	109-133	13	0.87
D13S1829 (193K19)	ATTTGGCTCCTATTTCCCTTTTGA GACACGCAATGAAGCCAGTGC	57	188-206	6	0.55
D13S1830 (748f2)[b]	TTTTGTTTGGTACCAGAAGGATTA CAGCCTGTGCGAGATAGTGAG	55	152-160	5	0.48
D13S1831 (35C18)	NCTGCCCACAAGAAACTCACTTCA TATTTTCCCCTTCCTTTCACTTTC	57	91-107	9	0.74
D13S1832 (63E10)	GATCGCTTGAGCCTAGGAGTTTG AAATGTATGGCTATGAGGTGTAGG	57	145-173	10	0.74
D13S1835 (114D9)	TCCATTTGCTCAGTAACTTTGTCC CTATATTCCAGCCTGGGCAAGAG	57	129-173	11	0.76

[a] 72H5 is a chromosome 21 PAC clone

[b] 748f2 is a YAC clone

identical segregation patterns were grouped together and considered as one marker in further analysis. One linkage group with a lod score threshold of 4.00 comprised 49 loci and was subsequently analyzed with the maximum likelihood multipoint program RHMAXLIK assuming two retention models, equal and centromeric. A stepwise locus ordering strategy that orders one locus at a time was used with a SAVMAX value set at 10-15 and a PRTMAX value at 3-5. With these parameters, all partial orders with maximum likelihoods within $10^{10\text{-}15}$ of the most likely partial order were calculated and only those within $10^{3\text{-}5}$ were printed. A framework map of a subset of loci ordered with odds greater than 1000:1 was produced and loci that could not be ordered with this level of support were added to the framework map one at a time. They were assigned by calculating maximum likelihood scores for all possible locations on this map and were positioned with a range of adjacent intervals. Distances were estimated in centirays, where 1 cR_{9000} corresponds to 1% breakage at an X-ray dosage of 9000 rads. We also used the list of locus orders option that allowed us to compare the likelihood for our order of a subset of loci to that of a YAC contig established in the region by Guilford *et al.* (1995).

Isolation of new chromosome 13q polymorphic markers

The pericentromeric chromosome 13q probes D13F263S2E, D13S1118, TUBA2, D13S1119, D13S183E, D13S1120, D13S1114, D13S1176, GJβ2, D13S1189, D13S502E, WI-8572 and D13S1117 were screened against the human genomic PAC library RPCI-1 at the CGAT Genome Resource Facility at the Hospital for Sick Children in Toronto and 57 PACs were isolated. PAC clones were grown overnight in TB broth supplemented

with 12.5 μg/ml kanamycin and DNA was extracted using a modified alkaline-lysis technique (Birnboim and Doly, 1979). To verify the chromosomal origin of these PACs, Southern blots containing *Eco*RI or *Hin*dIII-digested PAC DNA, total human DNA, DNA from a chromosome 13-only cell hybrid (GM11689), DNA from a chromosome 21-only cell hybrid (WAV17), and total hamster and mouse DNA, were prepared and hybridized against the probes sent originally to screen the PAC library. Southern blot analysis was done by standard methods (Sambrook *et al.*, 1989). Sau3A1 digests of PAC clones were subcloned into pUC18 *Bam*H1/BAP (Pharmacia Biotech). The resulting clones were transformed into DH5α competent cells (Life Technologies) and screened with a (polydA-dC). (polydG-dT) repeat probe (Pharmacia Biotech) by colony hybridization. Positive clones were identified and DNA was prepared from overnight cultures in TB broth supplemented with 60 μg/ml Ampicillin. The region surrounding the CA repeats was sequenced by the dideoxy chain termination method using a Sequenase Kit (USB). Suitable flanking PCR primer sequences were designed using Primer Select (DNASTAR).

In addition, four pericentromeric chromosome 13q YACs (987e10, 748f2, 911h8 and 967b1) were screened for the presence of $(CA)_n$ polymorphic repeats. YAC DNA was prepared by standard procedures as described elsewhere (Scherer and Tsui, 1991). Human-specific sequences were amplified from these clones by Alu-PCR using the Alu-specific primers TC-65, 517, 32, 33, 34, 515, 278 (Nelson *et al.*, 1989), A1S and CL (Ledbetter *et al.*, 1990). PCR conditions consisted of one denaturation cycle at 94°C for 4 minutes followed by 35 cycles each consisting of 1 min at 94°C and 45 sec. at 55°C followed by a final elongation cycle at 68°C for 5 min. The Alu-Alu PCR products were

separated on 2% agarose gels, blotted into Hybond-N^+ nylon membranes (Amersham) and hybridized against a (polydA-dC).(polydG-dT) repeat probe (Pharmacia Biotech). PCR products that gave positive signals were reamplified and subcloned into a plasmid vector pCR2.1 using the TA cloning kit (Invitrogen), as specified by the manufacturer and the resulting clones were screened for the presence of $(CA)_n$ repeats as described in the previous paragraph.

The new polymorphic markers were genotyped in 50 unrelated CEPH individuals as previously described (Kibar *et al.*, 1996) and the PIC value was calculated using the HET program of the LINKAGE 5.1 package (Terwilliger and Ott, 1994).

Fluorescence in situ hybridization (FISH) and Fiber-FISH

Preparation of Cells for Mapping and Extended Chromatin Fibers Release. The normal human lymphocytes were maintained and harvested in accordance with standard tissue culture protocols (Verma and Babu, 1995) and re-suspended in 1 to 2 ml PBS. Extended chromatin fibers were released according to a published method (Verma and Babu, 1995).

*FISH. Eco*RI-digested DNA products from PACs 15I2, 50F20, 187L17, 193K19 and 63E10, were purified using the GENECLEAN II kit (Bio/Can Scientific). These products were biotin-labeled or digoxigenin-labeled using a random-primed DNA labeling kit (Gibco-BRL, Gaithersburg, MD) and used as probes for FISH study. In situ hybridization (ISH) was performed by following Oncor's manual with a few modifications. For each slide, a 15µl mixture containing 100ng biotin-labeled DNA, 100ng digoxigenin-labeled DNA and 2µg human Cot-1 DNA (Gibco-BRL) in 70% formamide/2XSSC was

incubated in boiling water for 10 mins, and pre-annealed at 37°C water bath 40 minutes for suppression of repeating sequences. For Fiber-FISH analysis, digoxigenin-labeled chromosome 13/21 α satellite DNA probe (Oncor) was denatured at boiling water for 10 minutes and snap-chilled on ice just before hybridization. The extended free chromatin fiber and the interphase nuclei slides were dehydrated in an increasing grade (50%, 70%, 100%) ethanol series. Before hybridization, the slides were denatured in 70% formamide/2 x SSC, pH 7.0 at 74°C for 3 minutes, followed by rinsing in ice-cold 2 x SSC two times, 2 minutes each, and dehydration in an ice-cold ethanol series (50%, 70%, 100%), 3 minutes each. After hybridization in a humid chamber containing 50% formamide/2 x SSC at 37°C overnight, the slides were washed in 50% formamide/2 x SSC at 37°C twice, 7 minutes each and 2 x SSC at 37°C twice, 4 minutes each. Hybridization signals were detected with rabbit anti-biotin (Enzo Diagnostics, Farmingdale, NY) and mouse anti-digoxigenin (Boehringer Mannheim, Germany), 45 minutes each, and then with a mixture solution containing biotinylated goat anti-rabbit IgG (Gibco-BRL, Gaithersburg, MD) and sheep anti-mouse Ig-digoxigenin (Boehringer Mannheim) for 45 minutes, finally amplified with a mixture solution containing streptavidin-fluorescein isothiocyanate (FITC) conjugate (Gibco-BRL) and sheep anti-digoxigenin-rhodamine (Boehringer Mannheim) for 45 minutes. The slides were counterstained with 4,6-diamino-2-phenyl-indole (DAPI), mounted with antifade solution, and examined under a fluorescent microscope.

RESULTS

Two-point analysis of the RH data

Retention frequencies of the 54 loci included in this study range from 6.4 % (D13S860E, D13S851E, D13S1088E) to 65.6 % (D13Z1), with an overall average of 31.6 %. Two-point analysis allowed us to identify three linkage groups with a lod score threshold of 4.00: group 1 contains D13Z1 and 48 other loci shown in figure 4.2.(p. 93), group 2 contains D13F263S2E and D13S145 and group 3 contains D13S860E, D13S851E and D13S1088E. D13F263S2E is linked to D13Z1, TUBA2 and D13S1119 with a lod score threshold of 3.00 at distance estimates of 90 cR_{9000}, 94 cR_{9000} and 94 cR_{9000} respectively.

The 54 loci gave 43 unique retention patterns. Loci that gave identical retention patterns were: D13S1316-D13S141, D13S175-D13S1829, D13S1830-WI11739, WI8572-WI13846, D13S633-D13S250-D13S1125-D13S143-FB10F7-FGF9, D13F250S2E-D13S292, D13S851E-D13S1088E. These loci were treated as single mega-loci when included in the multipoint analysis.

Multipoint analysis of the RH data

We only considered linkage group 1 for multipoint analysis since it includes the centromeric marker D13Z1 and the distal flanking marker for the HED region D13S143. Maximum likelihood analysis was first carried out assuming an equal retention model. A 1000:1 framework map of 30 loci with 21 unique map positions was obtained, with a locus order 2932 more likely than the next order (data not shown). A decrease in locus

retention frequency was observed with increasing distance from the centromeric marker D13Z1 (figure 4.1.). Consequently, multipoint analysis was redone under a centromeric retention model. Only 3 loci out of 39 (D13Z1, D13S183E and D13S1835) could be ordered with 1000:1 level of support. D13Z1 represents alpha satellite DNA that consists of multiple copies of different higher-order tandem repeating structures of ~171 bp monomer units (Greig *et al.*, 1993). Hence, this marker doesn't detect a single locus and could complicate RH data analysis. When D13Z1 was excluded from the multipoint analysis, an equal retention model gave a 1000:1 framework map ordering 34 loci with 25 unique map positions, 1309 more likely than the next order while a centromeric retention model generated a 1000:1 framework map ordering 29 loci with 21 unique map positions, 1944 times more likely than the next order. The two maps have in common 24 loci with 16 unique map locations that have identical orders. When the 29 loci forming the centromeric framework map were reanalyzed under an equal retention model, the same best order was obtained and with different total map length and intermarker distance estimates (figure 4.3.). The centromeric model fit the data significantly better than the equal retention model ($\chi 2$=48.5, df = 1, $P < 0.00001$). Only the framework map obtained under the centromeric retention model and without D13Z1 is presented in figure 4.2. This map spans 347 cR_{9000} with an average intermarker distance of 17.3 cR_{9000} ranging from 2.6 cR_{9000} to 85.2 cR_{9000}. The other loci that could not be ordered with odds greater than 1000:1 were placed in a range of adjacent intervals.

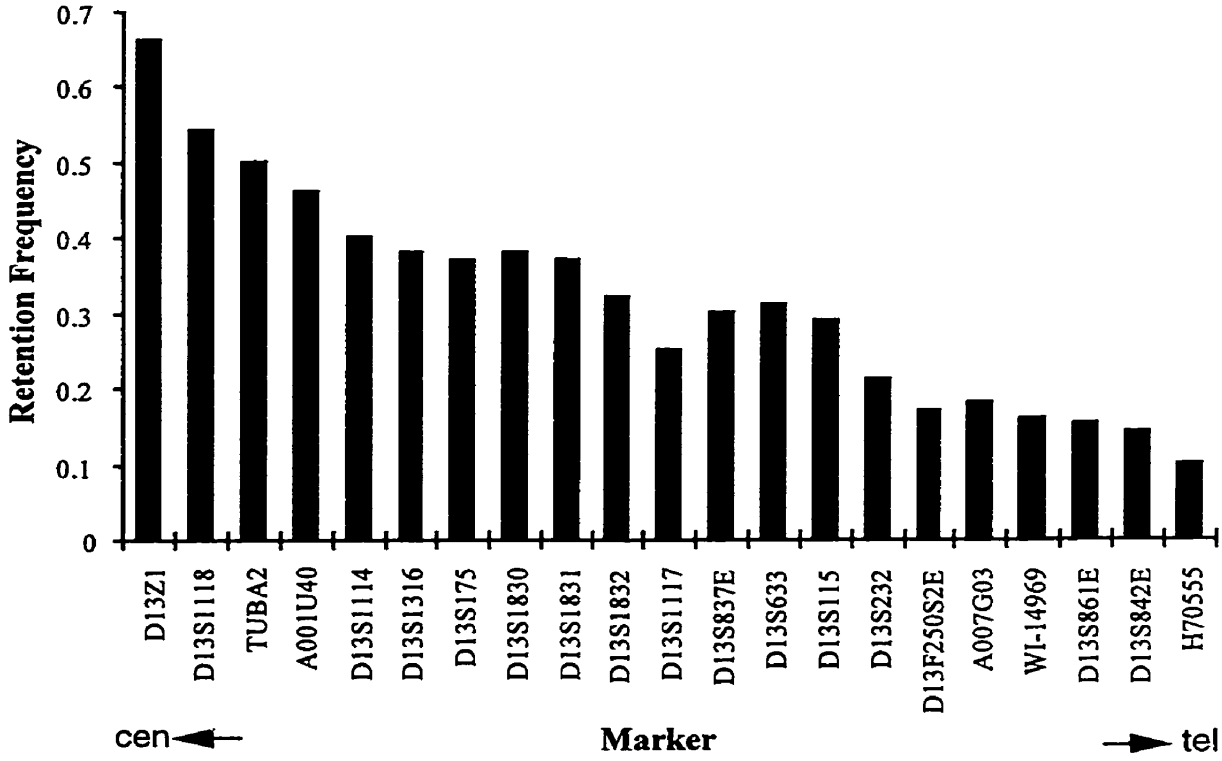

Figure 4. 1. Histogram of retention frequencies of 21 chromosome 13q pericentromeric markers in 94 radiation hybrid clones. The markers are arranged in the most likely 1000:1 order obtained under an equal retention model.

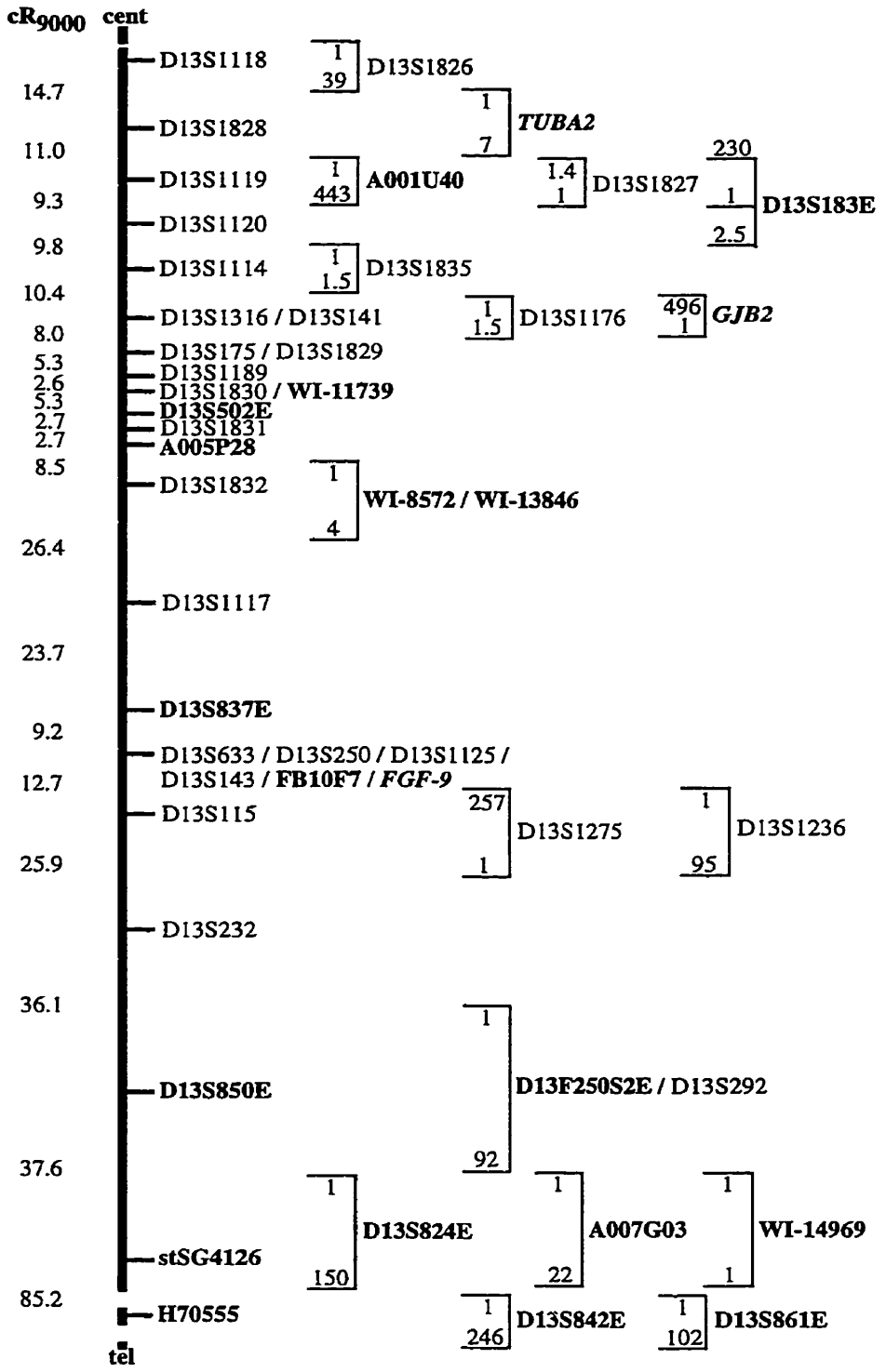

Figure 4. 2. 1000:1 framework radiation hybrid map of 29 loci spanning the HED candidate region on chromosome 13q. Estimated distances in cR_{9000} between adjacent loci are shown. Likelihood estimates for placement of the remaining 19 loci are shown on the right side of the figure. For a given marker, a value of 1 indicates the most likely interval and the other values indicate likelihood ratios against localization of that marker in each respective interval. Genes and ESTs are highlighted in italic and bold letters respectively.

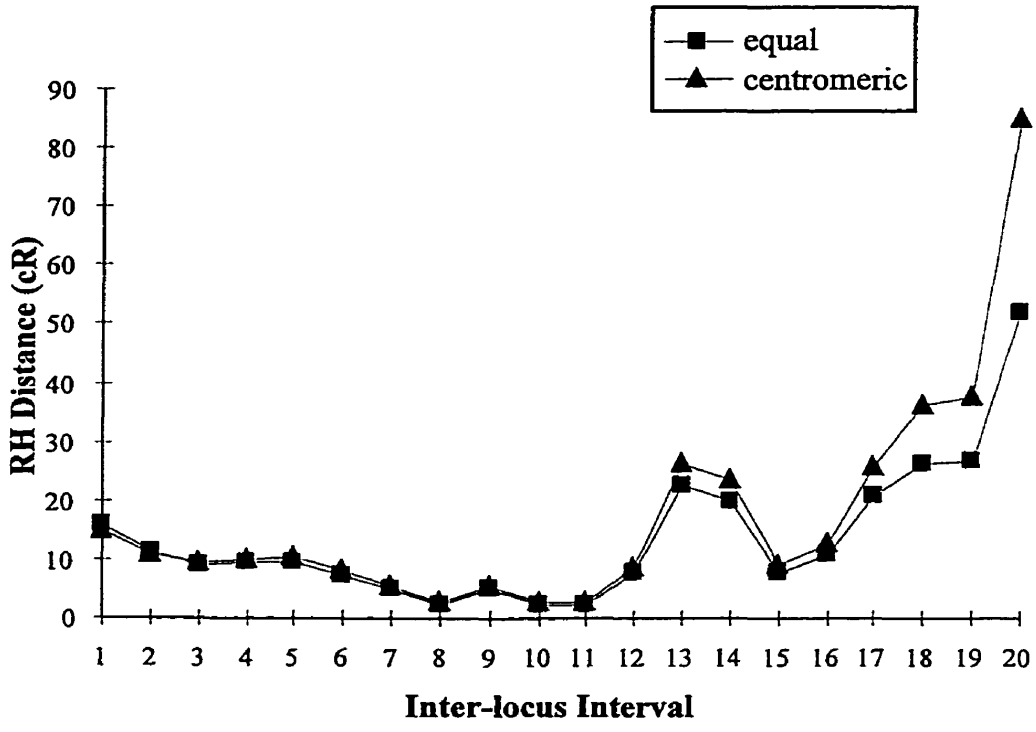

Figure 4. 3. Distance estimates in cR_{9000} for the 20 inter-loci intervals in the radiation hybrid map presented in figure 4. 2. These estimates were calculated under two different retention models: equal and centromeric.

New chromosome 13q polymorphic markers

After the RH map was constructed, 13 markers from the HED candidate region were screened against a human P1 genomic library and 57 PACs were isolated. While mapping these PACs by southern hybridization with the various loci, we encountered cross-hybridization with sequences other than that on chromosome 13 with three probes: D13F263S2E, D13S1118 and TUBA2. Extensive homology of the chromosome 13q and 21q centromeric regions has been reported and extends to unique pericentromeric sequences (Van Camp *et al.*, 1992). In fact, 4 out of 8 PACs isolated with D13S1118 are of chromosome 21 origin. The origin of the PACs isolated with the other 2 probes could not be determined because of the complex pattern of hybridization obtained perhaps because they are pseudogenes or members of gene families. Out of the 57 PACs obtained, only 22 clones were found to be of chromosome 13 origin.

Small insert libraries were generated for these 22 chromosome 13 PACs, 4 chromosome 21 PACs (isolated with chromosome 13q markers) and 4 YACs that partially cover the HED candidate region. Eight new polymorphic $(CA)_n$ markers were developed. The marker D13S1826 was isolated from a chromosome 21 PAC (72H5). The primers initially designed for this marker amplified from both chromosome 13-only and chromosome 21-only DNA. To design chromosome 13-specific primers, PCR products amplified from total human DNA, chromosome 13-only DNA and chromosome 21-only DNA were subcloned (as described for subcloning of Alu-Alu PCR products in Materials and Methods), several clones were picked and sequenced with new primers being designed with unique 3'ends. Table 4.1. (page 85) lists the isolated PACs or YACs, primer sequences, annealing temperature, PCR product size range, number of alleles and

PIC value for each of the 8 new markers. They have an average PIC value of 0.67, with 2 markers having PIC values > 0.80 and 3 markers > 0.70.

DISCUSSION

As a step towards the isolation of the HED gene, we have used a panel of 94 chromosome 13q RH clones to map 54 pericentromeric loci. The average retention frequency of these markers is 31.6 % with D13Z1 being the most highly retained (65.6 %). A remarkable gradient in retention frequencies was observed with markers closer to the centromere having increased retention frequencies compared to more distal ones (figure 1). The high retention frequency for centromeric alphoid repeats (D13Z1 in this study) and the "centromeric effect" have been reported in other radiation hybrid studies (Ceccherini *et al.*, 1992; James *et al.*, 1994; Francke *et al.*, 1994; Walter *et al.*; 1994) and for this panel in particular (Shaw *et al.*, 1995). This could be due to selective retention of a functional centromere or a section of a centromere that confers stability and growth advantage on chromosomal fragments or to selective retention against subtelomeric regions that have more expressed genes than pericentric regions and hence are less tolerated by the hamster recipient cell line.

Two-point analysis of RH data detected 3 linkage groups. Group 1 contained D13Z1, D13S143 and 47 other loci shown in figure2. Linkage group 2 contained D13S145 and D13F263S2E that is linked to D13Z1, TUBA2 and D13S1119 but at very large estimated distances. Linkage group 3 consisted of 3 loci (D13S860E, D13S851E and D13S1088E) and is most likely telomeric to the HED candidate region. Consequently, only group linkage 1 was further studied.

Maximum likelihood multipoint analysis was done under both equal and centromeric retention models with and without D13Z1. This centromeric marker could complicate RH analysis since it is highly retained and may detect more than one locus. An equal retention model gave a 1000:1 framework map of 30 loci with D13Z1 versus a 1000:1 framework map of 34 loci without D13Z1, with 24 loci common to both and ordered the same. Removal of D13Z1 from the analysis did affect the likelihoods of the order of non-adjacent distal loci (data not shown). On the other hand, a centromeric retention model with D13Z1 could not order more than 3 loci with odds $\geq$ 1000:1 while the same model without D13Z1 could generate a 1000:1 framework map of 29 loci with 21 unique map positions. This map, shown in figure 4.2., spans 347 cR_{9000} and has an order with maximum likelihood ~2000 times than any other order giving strong evidence that it is correct. The same best order was obtained for these 29 loci when the analysis was redone under an equal retention model with different map length (274 cR_{9000}) and intermarker distance estimates. As shown in figure 4.3., the difference in distance estimates calculated under the two different retention models becomes large at the telomeric side. These results are consistent with the computer simulation-based findings of Lunetta and Boehnke (1994) that the retention model assumed does not affect the best order inferred but could strongly influence the map length and intermarker distance estimates, especially with the presence of a strong centromeric effect.

Although RH mapping is a statistical method based on the assumption of random X-ray breakage along the chromosome, RH distances can closely reflect physical distances. Conversions of cR to kb are based on comparing either the distance separating two telomeric ends of a chromosome to its total DNA content (James *et al.*, 1994; Shaw

et al., 1995) or by developing STSs from the ends of YACs of known sizes and integrating these STSs into the RH map (Thompson *et al.*, 1993; Francke *et al.*, 1994) or from genetic mapping data (Abel *et al.*, 1992). Various estimates of the cR / kb ratio at an X-ray dosage of 9000 rads have been calculated including a ratio of 50 (Abel *et al.*, 1993; James *et al.*, 1994), of 55 (Richard *et al.*, 1991) and of 75 (Shaw *et al.*, 1995), albeit, from different panels. These different estimates could be due to many reasons including inaccurate estimation of either the chromosome DNA content or physical length, or the YACs size or the length of the total RH map depending on the conversion method used, use of different cell lines and chromosomal or regional sensitivity to irradiation. If we consider 1 cR_{9000} to approximate 50-75 kb, our RH map would have an average resolution of 17.3 cR_{9000} corresponding to ≅ 865 kb –1.3 Mb. The smallest interval is 2.6 cR_{9000} (130-195 kb) and the largest interval is 85.2 cR_{9000} (4.3-6.4 Mb).

A roughly linear relationship was found between genetic distance and RH mapping distance (Boehnke, 1992). Our RH data allows us to crudely estimate the ratio of cM to cR_{9000} in this region. According to our RH map, the intervals D13S141-D13S232 and D13S175-D13S232 span 133cR_{9000} and 125 cR_{9000} respectively. Genetic distances of these two intervals are estimated to be 4.2 cM (Warburton *et al.*, 1996) and 3.4 cM (Scott *et al.*, 1995) respectively. Comparing these distances would give an approximate ratio of 1 cM to 34 cR_{9000}. Consequently, if 1 cR_{9000} corresponds to 50-75 Kb, 1cM would approximate 1.7-2.5 Mb in this region. An increased Mb to cM ratio is expected in pericentromeric regions that undergo a low rate of recombination.

To determine consistency of order of our map, we compared it to a YAC contig and an EST map constructed by Guilford *et al.* (1995) in the region. A major discrepancy

in locus order was found in the interval D13S1118 – D13S175. The block D13S1118-TUBA2-D13S1119-D13S1120-D13S1114-D13S175 is inverted in the YAC contig map. Likelihoods for the two contested orders were compared and our RH order was found to be the best locus order with a log likelihood difference of 17 from the YAC contig order. To help resolve this inconsistency, FISH analysis was carried out with the PACs 15I2, 50F20, 187L17, 193K19 and 63E10 isolated with D13S1118, TUBA2, D13S1119, D13S175 and D13S1117 respectively. Fiber-FISH analysis with the PACs 15I2 and 193K19 demonstrated that D13S1118 is centromeric to D13S175. FISH analysis on interphase nuclei with the PACs 15I2, 50F20, 187L17 and 63E10 showed the orders D13S1118-TUBA2-D13S1117 and D13S1118-D13S1119-D13S1117. Figure 4.4. depicts the results obtained with FISH analysis using the PACs 15I2, 187L17 and 63E10. These results give support to our order and not to the YAC contig order. This inconsistency could be caused by YAC rearrangements such as chimerism and internal deletions and/or by the extensive homology reported between chromosomes 13 and 21 pericentric regions (Van Camp *et al.*, 1992). As mentioned before, 4 out of 8 PACs picked up with D13S1118 are of chromosome 21 origin indicating a high level of sequence identity between these two chromosomes in this region. Such sequence conservation could affect the ordering accuracy of the YAC contig especially that PCR results were not confirmed by Southern blot analysis.

In conclusion, we have constructed a radiation hybrid map of 48 chromosome 13q pericentromeric loci including the HED candidate region. This map integrates 3 genes (TUBA2, GJβ2 and FGF-9) and 18 ESTs with 27 other markers including 19 polymorphic markers. In addition, we have developed eight new polymorphic $(CA)_n$

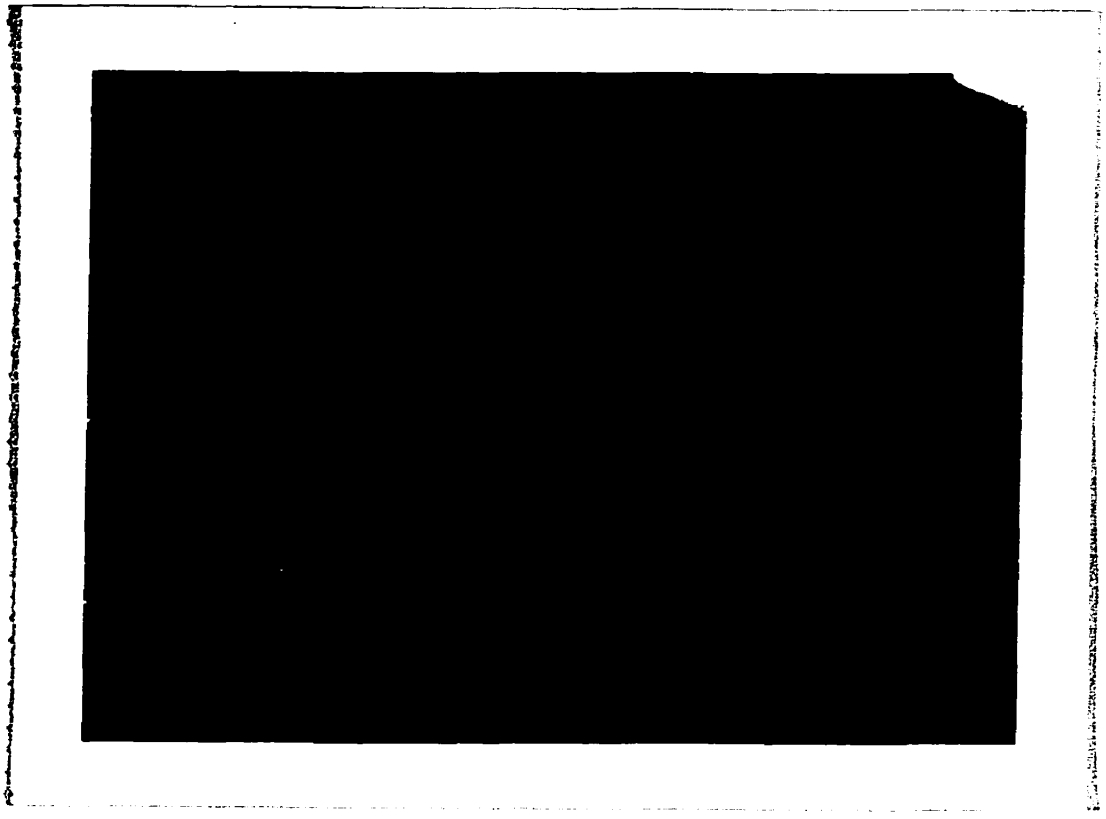

Figure 4.4. FISH analysis on interphase nuclei using the PACs 15I2, 187L17.and 63E10 isolated with the markers D3S1118, D13S1119 and D13S1117 respectively. Biotin-labeled probe (15I2) is indicated by an arrow and digoxigenin-labeled probes (187L17 and 63E10) are indicated by overheads. These results are consistent with the order D13S1118-D13S1119-D13S1117.

markers in this region, five of which are highly polymorphic. These resources will be useful for the isolation of the HED gene as well as other disease loci that are localized in the region.

ACKNOWLEDGMENTS

We would like to thank the families that actively participated in the study of HED. We are grateful to the CGAT Genome Resource Facility at the Hospital for Sick Children in Toronto for providing PAC clones. This work was supported by grants from the Canadian Dermatology Foundation and the National Foundation for Ectodermal Dysplasias (G.A.R.) and USNIH grant HD28088 (A.C.). G.A.R is supported by the Medical Research Council of Canada.

REFERENCES

Abel, K.J., Boehnke, M., Prahalad, M., Ho, P., Flejter, W. L., Watkins, M., VanderStoep, J., Chandrasekharappa, S. C., Collins, F. S., Glover, T. W., and Weber, B. L. (1993). A radiation hybrid map of the *BRCA1* region of chromosome 17q12-q21. *Genomics* **17:** 632-641.

Birnboim, H. C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acid Res.* **7:** 1513-1523.

Boehnke, M. (1992). Genetic analysis Workshop 7: radiation hybrid and somatic cell hybrid mapping of chromosome 21. *Cytogenet. Cell Genet.* **59:** 74-76.

Boehnke, M., Lange, K., and Cox, D. R. (1991). Statistical methods for multipoint radiation hybrid mapping. *Am. J. Hum. Genet.* **49:** 1174-1188.

Ceccherini, I., Romeo, G., Lawrence, S., Breuning, M. H., Harris, P. C., Himmelbauer, H., Frischauf, A. M., Sutherland, G. R., Germino, G. G., Reeders, S. T., and Morton, N. E. (1992). Construction of a map of chromosome 16 by using radiation hybrids. *Proc. Natl. Acad. Sci. USA* **89:** 104-108.

Clouston, H. R. (1929). Hereditary ectodermal dystrophy. *Can. Med. Ass. J.* **21**: 18-31.

Cox, D. R., Burmeister, M., Price, E. R., Kim, S., and Myers, R. M. (1990). Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. *Science* **250:** 245-250.

Der Kaloustian, V. M., and Kurban, A. K. (1979). Ecodermal dysplasias *In* "Genetic diseases of the skin" pp. 109-114, Springer-Verlag.

Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J., and Weissenbach, J. (1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380:** 152-154.

Escobar, V., Goldblatt, L. I., Bixler, D., and Weaver, D. (1983). Clouston sydrome: an ultrastructural study. *Clinical Genetics* **24:** 140-146.

Francke, U., Chang, E., Comeau, K., Geigl, E.-M., Giacalone, J., Li, X., Luna, J., Moon, A., Welch, S., and Wilgenbus, P. (1994). A radiation hybrid map of human chromosome 18. *Cytogenet. Cell Genet.* **66:** 196-213.

Goss, S. J., and Harris, H. (1975). New method for mapping genes in human chromosomes. *Nature* **255:** 680-684.

Greig, G. M., Warburton, P. E., and Willard, H. F. (1993). Organization and evolution of an alpha satellite DNA subset shared by human chromosome 13 and 21. *J. Mol. Evol.* **37:** 464-475.

Guilford, P., Dode, C., Crozet, F., Blanchard, S., Chaib, H., Levilliers, J., Levi-Acobas, F., Weil, D., Weissenbach, J., Cohen, D., Le Paslier, D., Kaplan, J-C., and Petit, C. (1995). A YAC contig and an EST map in the pericentromeric region of chromosome 13 surrounding the loci for neurosensory nonsydromic deafness (DFNB1 and DFNA3) and limb-girdle muscular dystrophy type 2C (LGMD2C). *Genomics* **29**: 163-169.

James, M. R., Richard III, C. W., Schott, J.-J., Yousry, C., Clark, K., Bell, J., Terwilliger, J. D., Hazan, J., Dubay, C., Vignal, A., Agrapart, M., Imai, T., Nakamura, Y., Polymeropoulos, M., Weissenbach, J., Cox, D. R., and Lathrop, G. M. (1994). A radiation hybrid map of 506 STS markers spanning human chromosome 11. *Nature Genet.* **8:** 70-76.

Kibar, Z., Der Kaloustian, V. M., Brais, B., Hani, V., Fraser, F. C., and Rouleau, G. A. (1996). The gene responsible for Clouston hidrotic ectodermal dysplasia maps to the pericentromeric region of chromosome 13q. *Hum. Mol. Genet.* **5:** 543-547.

Ledbetter, S. A., Nelson, D. L., Warren, S. T., and Ledbetter, D. H. (1990). Rapid isolation of DNA probes within specific chromosome regions by interspersed repetitive sequence polymerase chain reaction. *Genomics* **6:** 475-481.

Lunetta, K. L., and Boehnke, M. (1994). Multipoint radiation hybrid mapping: comparison of methods, sample size requirements, and optimal study characteristics. *Genomics* **21:** 92-103.

Lunetta, K. L., Boehnke, M., Lange, K., and Cox, D. R. (1996). Selected locus and multiple panel models for radiation hybrid mapping. *Am. J. Hum. Genet.* **59:** 717-725.

Nelson, D. L., Ledbetter, S. A., Corbo, L., Victoria, M. F., Ramirez-Solis, R., Webster, T. D., Ledbetter, D. H., and Caskey, C. T. (1989). *Alu* polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. *Proc. Natl. Acad. Sci. USA* **86:** 6686-6690.

Pertrukhin, K. E., Speer, M. C., Cayanis, E., Bonaldo, M. D. F., Tantravahi, U., Soares, M. B., Fischer, S. G., Warburton, D., Gilliam, T. C., and Ott, G. (1993). A microsattelite genetic linkage map of human chromosome 13. *Genomics* **15**: 76-85.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Analysis and cloning of eukaryotic genomic DNA. *In* "Molecular Cloning: A laboratory Manual" (N. Ford, C. Nolan, and M. Ferguson, Eds.), pp. 9.31-9.62, Cold Spring Harbor Laboratory Press.

Scherer, S. W., and Tsui, L.-C. (1991). *In* "Advanced Techniques in Chromosome research" (K. Adolph, Ed.), pp. 33-72, Marcel Dekker, Inc., New York.

Schuler, G. D., Boguski, M. S., Stewart, E. A., Stein, L. D., Gyapay, G., Rice, K., White, R. E., Rodriguez-Tome, P., Aggarwal, A., Bajorek, E., Bentolila, S., Birren, B. B., Butler, A., Castle, A. B., Chiannilkulchai, N., Chu, A., Clee, C., Cowles, S., Day, P. J., Dibling, T., Drouot, N., Dunham, I., Duprat, S., East, C., and Hudson, T. J. (1996). A gene map of the human genome. *Science* **274:** 540-546.

Scott, D. A., Carmi, R., Elbedour, K., Duyk, G. M., Stone, E. M., and Sheffield, V. C. (1995). Nonsyndromic autosomal recessive deafness is linked to the DFNB1 locus in a large inbred Bedouin family from Israel. *Am. J. Hum. Genet.* **57:** 965-968.

Shaw, S. H., Farr, J. E. W., Thiel, B. A., Matise, T. C., Weissenbach, J., Chakravarti, A., and Richard III, C. W. (1995). A radiation hybrid map of 95 STSs spanning human chromosome 13q. *Genomics* **27**: 502-510.

Still, I. H., Roberts, T., Bia, B., Hawthorn, L., Auffray, C., and Cowell, J. (1996). Incorporation of 35 novel gene transcripts into the physical and genetic map of human chromosome 13. *Genomics* **33**: 159-166.

Terwilliger, J. D., and Ott, J. (1994). *In* "Handbook of Human Genetic Linkage", pp. 277-278, the John Hopkins University Press, London.

Thompson, T. G., Morrison, K. E., Kleyn, P., Bengtsson, U., Gilliam, T. C., Davies, K. E., Wasmuth, J. J., and McPherson, J. D. (1993). High resolution physical map of the region surrounding the spinal muscular atrophy gene. *Hum. Mol. Genet.* **2**: 1169-1176.

Van Camp, G., Cruts, M., Backhovens, H., Wehnert, A., and Van Broeckhoven, C. (1992). Unique sequence homology in the pericentromeric regions of the long arms of chromosomes 13 and 21. *Genomics* **12**: 158-160.

Verma, R. S., and Babu, A. (1995). *In* "Human chromosomes: Principles and Techniques", pp. 9-11, 185-187, McGraw-Hill, New York.

Walter, M. A., Spillett, D. J., Thomas, P., Weissenbach, J., and Goodfellow, P. N. (1994). A method for constructing radiation hybrid maps of whole genomes. *Nature Genet.* **7**: 22-28.

Warburton, D., Shaw, S. H., Matise, T. C., Kalachikov, S., and Fischer, S. (1996). Report of the third international workshop on human chromosome 13 mapping 1995. *Cytogenet. Cell Genet.* **75**: 85-110.

Williams, M., and Fraser, F. C. (1967). Hydrotic ectodermal dysplasia-Clouston's family revisited. *Can. Med. Ass. J.* **96**: 36-38.

CHAPTER 5

Clouston hidrotic ectodermal dysplasia (HED): genetic homogeneity, presence of a founder effect in the French Canadian population and fine genetic mapping

Isolation of new polymorphic markers and collection of additional HED families allowed to study genetic homogeneity in HED and the presence of a founder effect in the French Canadian population and to further narrow down the HED candidate region

The presented work was submitted as:

Kibar, Z., Dubé, M. –P., Powell, J., McCuaïg, C., Hayflick, S. J., Zonana, J., Hovnanian, A., Radhakrishna, U., Antonarakis, S. E., Benohanian, A., Sheeran, A. D., Stephan, M. L., Gosselin, R., Kelsell, D. P., Christianson, A. L., Fraser, F. C., Der Kaloustian, V. M., and Rouleau, G. A. Clouston hidrotic ectodermal dysplasia (HED): genetic homogeneity, presence of a founder effect in the French Canadian population and fine genetic mapping.

In this chapter, references were reformatted and listed in alphabetical sequence by author's last name to be consistent with other chapters in the thesis

Contribution by co-authors:

Kibar: DNA isolation and genotyping; CRI-MAP analysis; two-point linkage analysis; haplotype construction, study of allelic association

Dubé: multipoint linkage and linkage disequilibrium analyses

Powell, McCuaïg, Hayflick, Zonana, Hovnanian, Radhakrishna, Antonarakis, Benohanian, Sheeran, Stephan, Gosselin, Kelsell, Christianson, Fraser and Der Kaloustian: collection of HED family material

Rouleau: lab. supervisor

ABSTRACT

HED is an autosomal dominant skin disorder that is particularly common in the French Canadian population. We previously mapped the HED gene to the pericentromeric region of chromosome 13q using linkage analysis in eight French Canadian families. In this study, we extend our genetic analysis to include a multiethnic group of 29 families with 10 polymorphic markers spanning 5.1 cM in the candidate region. Two-point linkage analysis strongly suggests absence of genetic heterogeneity in HED in four families of French, Spanish, African and Malaysian origins. Multipoint linkage analysis in all 29 families generated a peak Lod score of 52.13 at D13S1835 with a 1-lod-unit support interval spanning 1.9 cM. Recombination mapping placed the HED gene in a 2.4 cM region flanked by D13S1828 proximally and D13S1830 distally. We next show evidence for a strong founder effect in families of French Canadian origin and use linkage disequilibrium mapping to further refine the localization of the HED gene. Significant association was found between HED in these families and all markers analyzed (Fisher's exact test, $P<0.001$). Complete allelic association was detected at D13S1828, D13S1827, D13S1835, D13S141 and D13S175 ($P_{excess} = 1$) spanning 1.3 cM. A major haplotype including all 10 associated alleles was present on 65% of affected chromosomes. This haplotype most likely represents the founder haplotype that introduced the HED mutation into the French Canadian population. Multipoint linkage disequilibrium analysis based on the likelihood approach developed by J. D. Terwilliger (1995) estimates the most likely position of the HED gene at 0.08 cM telomeric to D13S1835, with a 3.27 cM calculated support interval.

INTRODUCTION

Clouston hidrotic ectodermal dysplasia (HED) is a rare form of genodermatosis transmitted as an autosomal dominant condition with complete penetrance and variable expressivity (Williams and Fraser, 1967). Though HED has been described in families of various ethnic origins (McNaughton and Pierson, 1976; Rajagopalan and Hai Tai, 1977; Ando *et al.*, 1988; Patel *et al.*, 1991; Radhakrishma *et al.*, 1997; Taylor *et al.*, 1998), it is particularly common in families of French Canadian origin suggesting a founder effect for the HED mutation in this population (Clouston 1929; 1939; Kibar *et al.*, 1996). The main features of HED are nail hypoplasia and deformities frequently associated with paronychial infections, hair defects that range from brittleness and slow growth rate to partial or total alopecia and palmoplantar hyperkeratosis that can vary from moderate to severe and diffuse thickening of the skin (Der Kaloustian and Kurban, 1979). Less frequently, familial cases of HED have been described with other anomalies including hyperpigmentation of the skin, thickening of the skull bones, tufting of the terminal phalanges, mental deficiency, hearing loss and ocular abnormalities (Pierard *et al.*, 1979, Worobec-Victor *et al.*, 1988).

Biochemical and biophysical studies of the defective hair and nails in HED suggest a molecular abnormality of the hard keratins or their associated proteins (Gold and Scriver, 1972; Giraud *et al.*, 1977; Escobar *et al.*, 1983). Using linkage analysis in eight French Canadian families segregating HED, we initially excluded linkage of HED to these candidate genes and mapped the HED gene to the pericentromeric region of chromosome 13q. The highest two-point lod score obtained in that study was 8.12 at zero recombination from the marker D13S175. A multipoint linkage analysis using the

markers D13S175, D13S141 and D13S143 gave a maximum lod score of 11.12 at D13S141 with the one-lod-unit support interval spanning a 12.7 cM region which includes D13S175 and D13S141. Previous haplotype analysis using the markers cent-D13S175-D13S141-D13S143-D13S115-D13S1236-D13S292-D13S283-tel allowed us to place the HED gene in a region spanning at least 15 cM between the centromere and D13S143 (Kibar *et al.*, 1996). Later studies have confirmed linkage of HED to this region in an ethnically diverse group of families (Indian, French and Scottish-Irish), demonstrating absence of genetic heterogeneity in HED (Radhakrishma *et al.*, 1997; Taylor *et al.*, 1998). However, these studies failed to detect recombinants that would narrow down our established candidate region.

As a step towards the identification of the HED gene, we constructed a radiation hybrid map of 48 loci on chromosome 13q including the HED locus and isolated eight new polymorphic markers in the candidate region (Kibar *et al.*, 1999). In the current study, we genotype an extended panel of 29 HED families of various ethnic origins with a set of 10 polymorphic markers to study genetic homogeneity in HED and finely map the HED gene. In addition, we present evidence for a founder effect for this disease in the French Canadian population and use linkage disequilibrium mapping to further refine the HED-containing interval.

RESULTS

Two-point linkage analysis in the non French Canadian families

Six families of various ethnic origins- Indian, Scottish-Irish, French, Spanish, South African and Malaysian- were tested for linkage of the HED gene to 10

pericentromeric chromosome 13q markers. Linkage in the Indian and Scottish-Irish families to this region was previously demonstrated (Radhakrishma *et al.*, 1997; Taylor *et al.*, 1998); this study extends the genotyping analysis of these families with 7 new markers. Table 5.1 shows the maximum lod score and the maximum recombination fraction obtained with each marker. Most maximum lod scores were obtained at zero recombination from the markers analyzed. Lod scores above 3.00 were obtained only with the Indian and Scottish-Irish families. However, the Malaysian family gave a maximum lod score very close to 3.00 ($Z_{max} = 2.93$) with the marker D13S1835. The South African, Spanish and French families generated maximum lod scores of 2.11, 1.20 and 1.80 respectively. Since these lod scores are equal to the maximum possible lod scores those families can generate assuming 100% informativity of markers, the HED gene in these families is most probably linked to the markers analyzed. Lod scores <-2.00 were obtained only in the French family with the markers D13S1832 and D13S143 indicating the presence of a recombination event with the two markers. These results suggest absence of genetic heterogeneity in HED in the multiethnic group of families analyzed.

Multipoint linkage analysis

A multipoint linkage analysis was carried out to use information in all 29 HED families. In each test run, we were able to analyze only four loci simultaneously. After covering the entire map, we repeated the analysis with the three most informative markers chosen from the runs that generated the highest lod scores. Results of multipoint linkage analysis using the genetic map: D13S1828-0.51 cM-D13S1835-0.81 cM-D13S175 are demonstrated in figure 5.1. A maximum lod score of 52.13 was obtained at D13S1835.

Table 5. 1. Absence of genetic heterogeneity in HED

Family origin	Maximum Lod score at recombination fraction = 0 from marker									
	S1826	S1828	S1827	S1835	S141	S175	S1830	S1831	S1832	S143
Indian	7.97	6.94	5.63	0	1.57	1.28[a]	6.09	7.60	5.70	0.38
Scottish-Irish	1.11	3.44	0	3.17	2.88	1.93	3.55	1.72	3.61	1.11
French	1.78	1.79	1.79	1.80	1.02	0.75	0.24	0.19	-4.34×10^{19}	-4.34×10^{19}
Spanish	1.20	1.20	0.90	0.44	1.20	0.90	0	0.90	0.60	1.20
South African	2.11	0.91[b]	-0.04	2.11	-0.10	0.76	-0.05	1.81	1.81	-0.03
Malaysian	2.04	2.67	2.81	2.93	0.38	1.38	0.10[c]	1.05	0.35	0.59

Recombination fractions: [a]0.035, [b]0.117 and [c]0.111.

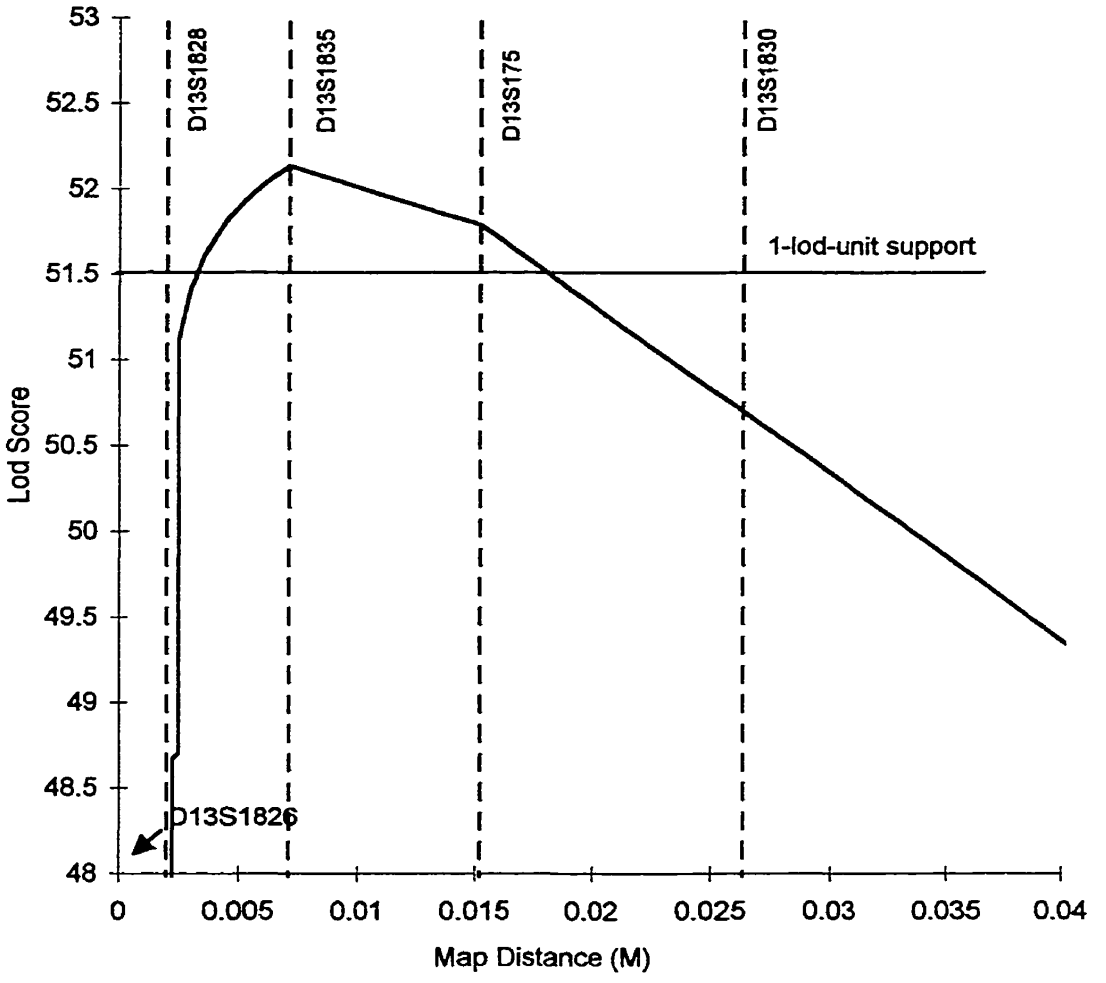

Figure 5. 1. Multipoint linkage analysis of 29 HED families using the genetic map: D13S1828- 0.51 cM-D13S1835-0.81 cM-D13S175.

The one-lod-unit support interval spans 1.93 cM, 0.05 cM telomeric to D13S1828 and 0.43 cM centromeric to D13S1830.

Recombination mapping

Haplotypes were constructed manually, assuming a minimum number of recombination events. Nine observed recombinants were detected in our families (data not shown), three of which have been previously described (Kibar *et al.*, 1996). Figure 5.2. shows the two most critical recombinants that would define the new boundaries for the HED candidate region: individuals II-5 in family A and II-4 in family B that would place the HED gene telomeric to D13S1828 and centromeric to D13S1830 respectively. The region flanked by these two markers spans 2.4 cM. Both families are of French Canadian origin and share the same haplotype except at D13S1831, D13S1832 and D13S143 indicating the presence of a historical recombinant at D13S1831. No double recombinants were identified.

Allelic association and haplotype analysis

We analyzed 20 apparently unrelated French Canadian families for allelic association between HED and 10 markers in the candidate region. The results are shown in table 5.2. Significant evidence for linkage disequilibrium was found at all markers studied ($P<0.001$). The strongest allelic association was detected between HED and alleles at markers D13S1828, D13S1827, D13S1835, D13S141 and D13S175. These markers had a P_{excess} value of 1 indicating the presence of the associated alleles in 100% of affected chromosomes. While alleles 4 at D13S1827 and 2 at D13S141 are the most common

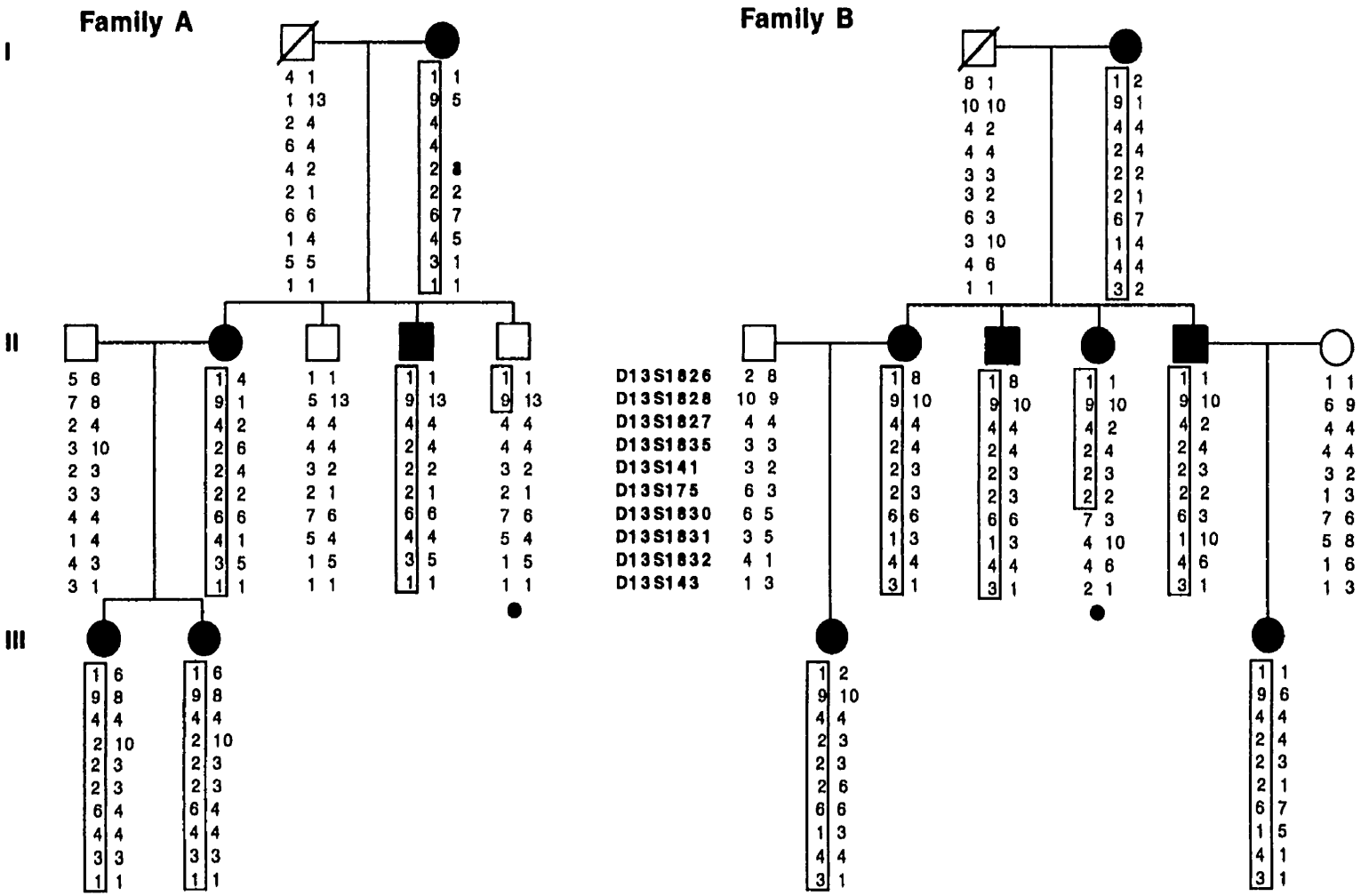

Figure 5. 2. Pedigrees of two HED families showing two recombinants that would place the HED gene telomeric to D13S1828 and centromeric to D13S1830. Phasing of haplotypes was deduced from available genotypes. The affected haplotypes are boxed. The order of markers used was obtained from a RH map constructed in the region (Kibar *et al.*, in press) and analysed with CRI-MAP using combined data from 14 CEPH families and 20 HED families.

Table 5. 2. Allelic association between HED and chromosome 13q markers

Marker	Disease-associated	Allele frequency[a]		P^b	$P_{excess}{}^c$
	allele (size in bp)	Affected	Normal		
D13S1826	1 (112)	0.92	0.33	1.9×10^{-7}	0.885
D13S1828	9 (127)	1.00	0.16	$< 10^{-9}$	1.000
D13S1827	4 (189)	1.00	0.67	2.9×10^{-4}	1.000
D13S1835	2 (129)	1.00	0.00	$< 10^{-9}$	1.000
D13S141	2 (124)	1.00	0.57	1.1×10^{-5}	1.000
D13S175	2 (104)	1.00	0.18	$< 10^{-9}$	1.000
D13S1830	6 (158)	0.96	0.50	1.5×10^{-5}	0.924
D13S1831	1 (91)	0.85	0.14	$<10^{-9}$	0.820
D13S1832	4 (163)	0.96	0.48	1.4×10^{-5}	0.923
D13S143	3 (128)	0.73	0.15	2.4×10^{-6}	0.682

[a] The number of chromosomes studied was 26 affected and 56-58 normal.

[b] The *P* value was generated using Fisher's exact test on a 2 x 2 table.

[c] $P_{excess} = (P_{affected} - P_{normal})/(1 - P_{normal})$

alleles on the normal chromosomes (67% and 57% respectively), alleles 9 at D13S1828 and 2 at D13S175 are less common (16% and 18% respectively). Allele 2 at D13S1835 is not present on any control chromosome presenting an almost complete disequilibrium. The allelic excess drops to 0.68 with the most telomeric marker D13S143.

The significant allelic association between HED and the markers analyzed strongly supports the founder effect hypothesis for the HED mutation in the French Canadian population. We consequently examined the affected chromosomes for the presence of an ancestral haplotype that most likely carried the founding mutation and of historical recombinants that would help us map the HED gene. The haplotypes present on the affected French Canadian chromosomes are shown in table 5.3. One major ancestral haplotype 1-9-4-2-2-2-6-1-4-3 was present on 17/26 (65%) of the affected chromosomes. In addition, six minor haplotypes that most likely represent historical recombinants were detected. These recombinants define a core haplotype 9-4-2-2-2 at markers D13S1828, D13S1827, D13S1835, D13S175 and D13S141 that was shared by all affected chromosomes and was present on none of the control chromosomes. This core haplotype defines a 1.3-2.6 cM region that most likely harbors the HED gene.

Multipoint linkage disequilibrium analysis

To further refine the position of the HED gene, a multipoint linkage disequilibrium analysis was performed using the likelihood method developed by J. D. Terwilliger (1995). This method is a likelihood-based approach that uses all alleles at each marker locus without assumption of the associated allele and then combines pairwise likelihoods by multiplication for larger number of loci. Results of this analysis

Table 5. 3. HED-Affected haplotypes in 20 French Canadian families. The common haplotype is shaded.

No of affected chromosomes	Markers S1826	S1828	S1827	S1835	S141	S175	S1830	S1831	S1832	S143
17	1	9	4	2	2	2	6	1	4	3
1	11	9	4	2	2	2	6	1	4	3
1	9	9	4	2	2	2	6	1	4	3
3	1	9	4	2	2	2	6	1	4	1
2	1	9	4	2	2	2	6	10	4	1
1	1	9	4	2	2	2	6	4	3	1
1	1	9	4	2	2	2	7	4	4	2

are depicted in figure 5.3. The maximum likelihood estimate was found to be 398.84 located 0.08 cM telomeric to D13S1835 and 0.72 cM centromeric to D13S141. The calculated 13.8-unit-support interval spans an interval of 3.27 cM, 1.11 cM centromeric to D13S1826 and 0.66 cM telomeric to D13S141. The heterogeneity parameter α of the DISMULT routine was 1 for the observed maximum likelihood, and the decay parameter 6.

DISCUSSION

Genetic homogeneity in HED was previously demonstrated in families of French Canadian, Indian and Scottish-Irish descent (Kibar *et al.*, 1996; Radhakrishma *et al.*, 1997; Taylor *et al.*, 1998). Our results extend this finding to include families of African, Spanish, French and Malaysian origins. The African family was diagnosed as a separate entity from HED based mainly on the associated abnormal cardiac findings and absence of palmoplantar hyperkeratosis, skin hyperpigmentation, and skeletal anomalies in the affected members (Christianson and Fourie, 1996). However, our two-point analysis in this family suggests its linkage to the HED candidate region with a maximum lod score of 2.11 at zero recombination from D13S1826 and D13S1835. Consequently, we postulate that this family segregates either a new allele of the HED gene or one mutated gene responsible for this form of ectodermal dysplasia and closely linked to the HED gene or the HED gene with another defective gene responsible for the cardiac findings. As shown in table 5.4., a comparison of the affected chromosomes in the 29 families revealed seven different haplotypes. These kindreds most likely carry distinct mutations of the HED gene, suggesting the presence of numerous ancestral mutations.

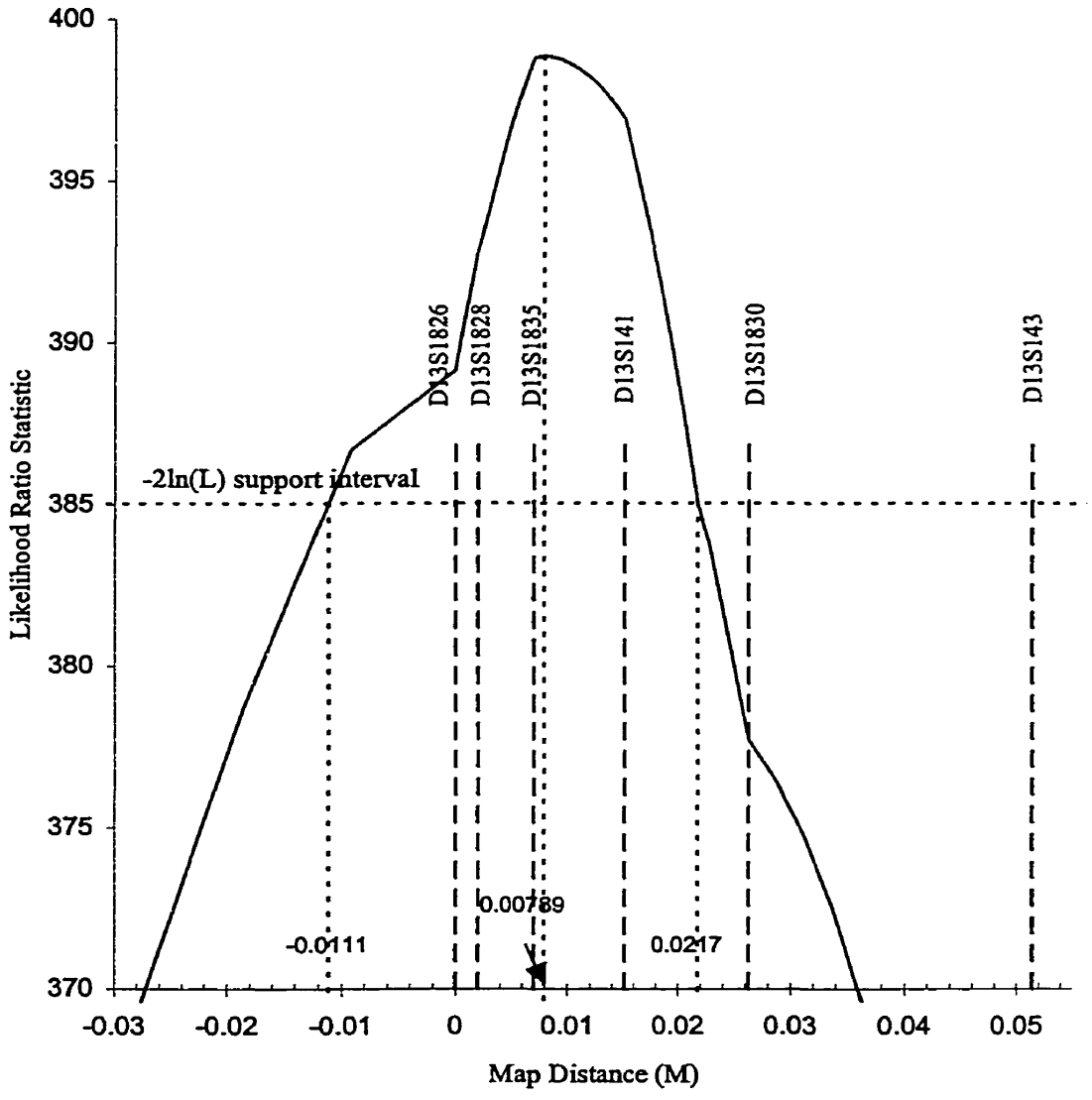

Figure 5.3. Multipoint linkage disequilibrium analysis of 20 French Canadian families using the DISMULT version 2.1 of the likelihood method developed by J. D. Terwilliger (1995). The genetic map used was: D13S1826-0.2-D13S1828-0.5-D13S1827-0.01-D13S1835-0.8-D13S141-0.01-D13S175-1.1-D13S1830-0.01-D13S1831-0.01-D13S1832-2.5-D13S143.

Table 5. 4. HED- Affected haplotypes of seven different origins

Family origin	Markers (allele size in bp)									
	S1826	S1828	S1827	S1835	S141	S175	S1830	S1831	S1832	S143
Indian	4 (118)	6 (121)	3 (187)	ND[a]	2 (124)	3 (106)	3 (152)	2 (95)	6 (167)	1 (124)
Scottish-Irish	9 (128)	5 (119)	4 (189)	11 (161)	2 (124)	5 (110)	3 (152)	5 (101)	11 (151)	1 (124)
French	9 (128)	5 (119)	5 (191)	12 (169)	3 (126)	3 (106)	7 (160)	5 (101)	4 (163)	2 (126)
French Canadian	1 (112)	9 (127)	4 (189)	2 (129)	2 (124)	2 (104)	6 (158)	1 (91)	4 (163)	3 (128)
Spanish	7 (124)	10 (129)	4 (189)	10 (159)	1 (114)	3 (106)	6 (158)	9 (109)	3 (161)	1 (124)
African	2 (114)	5 (119)	4 (189)	5 (135)	2 (124)	3 (106)	6 (158)	3 (97)	8 (171)	1 (124)
Malaysian	6 (122)	6 (121)	5 (191)	2 (129)	2 (124)	2 (104)	6 (158)	6 (103)	2 (159)	1 (124)

[a] No Data

Our initial linkage report has mapped the HED gene to a region extending from the centromere to D13S143 on chromosome 13q. This region spanning at least 15 cM is too large for positional cloning efforts (Kibar *et al.*, 1996). In this study, we used genetic analysis of 29 HED families with 10 polymorphic markers to narrow down the candidate interval. Multipoint linkage analysis generated a maximum lod score of 52.13 at D13S1835 with a 1-lod-unit support interval spanning 1.9 cM and including the markers D13S1827, D13S1835, D13S141 and D13S175. Recombination mapping placed the HED gene in a 2.4 cM region flanked by D13S1828 at the centromeric side and D13S1830 at the telomeric one.

The high frequency of HED in the French Canadian population and the particular population history suggests the presence of a founder effect. We tested our 20 unrelated French Canadian families for allelic association at 10 markers isolated from the candidate region and spanning a region of 5.1 cM. We found significant evidence for linkage disequilibrium at all of these markers with the strongest association at markers D13S1828, D13S1827, D13S1835, D13S141 and D13S175 ($P_{excess} = 1$). In addition, 65% of affected chromosomes shared the same haplotype that is most likely the ancestral one carrying the conserved HED mutation. These findings support the hypothesis of a major contribution from one founding HED mutation in the French Canadian population.

The French Canadian population of Quebec was founded by 8,000-10,000 colonists who came mostly from the central and western provinces of France during the period 1608-1759. These early settlers established themselves along the St. Lawrence River and, by the mid-19th century, many of them migrated towards the United States and to different regions of Quebec (Boleda, 1984; Desjardins, 1990; Heyer, 1995). Clouston

who first studied the hereditary nature of HED in a large French Canadian family living in Huntingdon in the Southwestern part of Québec described the remarkably high frequency of this disease in this part of the province (Clouston, 1929; 1939). In fact, most of our families originate from that region. Clouston inferred that the mutation was introduced into the French Canadian population by a French individual some time prior to the year 1759. Consequently, the French Canadian haplotype carrying the HED mutation most likely has a French origin. We have samples from one French family originating from Lorraine in the eastern part of France. The affected haplotype in this family shares only allele 4 at D13S1832 with the French Canadian haplotype. However, this marker resides outside the HED candidate region and hence the French haplotype most likely carries a historically distinct mutation. There is strong evidence for a founder effect in many rare French Canadian genetic diseases including oculopharyngeal muscular dystrophy (Brais *et al.*, in press), peripheral neuropathy with or without agenesis of the corpus callosum (Casuabon *et al.*, 1996), pseudo-vitamin D-deficiency rickets (Labuda *et al.*,.1996) and many others (reviewed in Bouchard and De Braekeleer, 1991). While most of these diseases originated around 12 generations ago and are particularly frequent in Northeastern Quebec, HED is believed to have been introduced later into the French Canadian population (8-9 generations ago) and is most frequent in the southwestern part of the province.

The presence of a founder effect for HED in the French Canadian population makes linkage disequilibrium (LD) mapping a possible tool to refine the localization of the HED gene. The principle of LD mapping is that in founder populations affected chromosomes descending from the same ancestral chromosome should share alleles at

polymorphic markers in the vicinity of the mutation. The extent of LD depends on the recombination fraction and hence genetic distances between the disease mutation and polymorphic markers, and on the demographic factors of the population such as generation depth, population growth and admixture. Other factors, notably mutation, selection and genetic drift, can also affect the pattern of LD (Devlin and Risch, 1995; Jorde, 1995). Although these factors have made some investigators cautious in using LD mapping results (Hill and Weir, 1994), various methods have been developed to incorporate their effect (summarized in Devlin *et al.*, 1996; Xiong and Guo, 1997) and LD mapping has been successfully used in fine-scale mapping of disease mutations in many founder populations such as the Finnish and French Canadian populations (Brais *et al.*, in press; Labuda *et al.*, 1996; De La Chapelle and Wright, 1998). Simple LD mapping that relies on the extent of disequilibrium placed the HED gene in a 1.3 –2.6 cM region including the markers D13S1828, D13S1827, D13S1835, D13S141 and D13S175. All affected chromosomes shared the same alleles at these markers giving rise to a core haplotype not present on any control chromosome. Consequently, we applied a multipoint LD method developed by J. D. Terwilliger (1995) to estimate the most likely position of the HED gene. Our LD multipoint analysis places the HED gene most likely at 0.08 cM telomeric to D13S1835 and 0.72 cM centromeric from D13S141 with a support interval spanning 3.27 cM.

Mutations at microsatellite loci can greatly affect the pattern of LD in the vicinity of the disease locus. Allele 7 (160 bp) at D13S1830 that defines the telomeric boundary of our new candidate region could be the result of slippage of the most common allele 6 (158 bp). However, this possibility would not affect the mapping results significantly

since the next telomeric marker D13S1831 that would be the new flanking marker is very close to D13S1830 (0.01 cM).

We can crudely estimate the ratio cM to Mb in the HED candidate region. Comparisons of our RH map to two genetic maps (Scott *et al.*, 1995; Warburton *et al.*, 1996) gives a ratio of 1 cM to 35 cR_{9000}. Consequently, if 1 cR_{9000} corresponds to 50 Kb, 1 cM would approximate 1.8 Mb in this region. An increased Mb to cM ratio is expected in pericentromeric regions that undergo a low rate of meiotic recombination. If this estimate were correct, our recombination mapping would place the HED gene in a 4.3 Mb (2.4 cM) region which is still large for positional cloning efforts. However, both multipoint analyses (linkage and LD) estimate the most likely position of the HED gene close to D13S1835, at 0 Kb from and 144 Kb telomeric to this marker respectively. If we overlap the support intervals obtained by both analyses, we will obtain a 3.4 Mb region (1.9 cM), 830 Kb centromeric to D13S1827 and 1.17 Mb telomeric to D13S175. Most likely, based on the LD analysis, the HED gene resides in the centromeric part of this region. These results will guide our strategies for physical mapping and cloning of the HED gene and we will aim at constructing a physical map around D13S1827, D13S1835, D13S141 and D13S175. Finer mapping of the HED gene by genetic analysis requires the collection of additional families and isolation of new polymorphic markers from the candidate region. However, it will be difficult to find more recombinants in this region because of its pericentromeric localization.

SUBJECTS, MATERIALS AND METHODS

HED Families

Twenty-nine families segregating HED were available for this study: 23 French Canadian, one Indian, one Scottish-Irish, one French, one Spanish, one South African and one Malaysian. They include 172 affected among a total of 287 subjects. The Indian, Scottish-Irish and South African families were previously described (Christianson and Fourie, 1996; Hayflick *et al.*, 1996; Radhakrishma *et al.*, 1997). The French family was collected by A. H. in the U.K.; two branches of this family have been reported by Giraud *et al.* (1977) and Taylor *et al.* (1998). The Malaysian family was described by Stevens *et al.* (in press). The Spanish family was collected from the United States. Pedigrees of the French, Spanish, African and Malaysian families are shown in figure 5.4. The French Canadian sample included the eight families we described initially in our linkage report, two families described by Taylor *et al.* (1998) and 13 new families. All patients presented with various degrees of nail defects, hair abnormalities and palmoplantar hyperkeratosis. In addition, the African family had abnormal cardiac findings associated with HED.

DNA and Microsatellite Analyses

DNA extraction and genotyping were done as previously described (Kibar *et al.*, 1996). Ten $(CA)_n$ polymorphic markers were used to genotype the HED families. They include 7 new markers we have recently isolated in the candidate region, D13S1826, D13S1828, D13S1827, D13S1835, D13S1830, D13S1831 and D13S1832 (Kibar *et al.*, 1999) along with D13S175, D13S141 and D13S143. The order of these markers was obtained from the RH map we constructed in the region with 1000:1 level of significance.

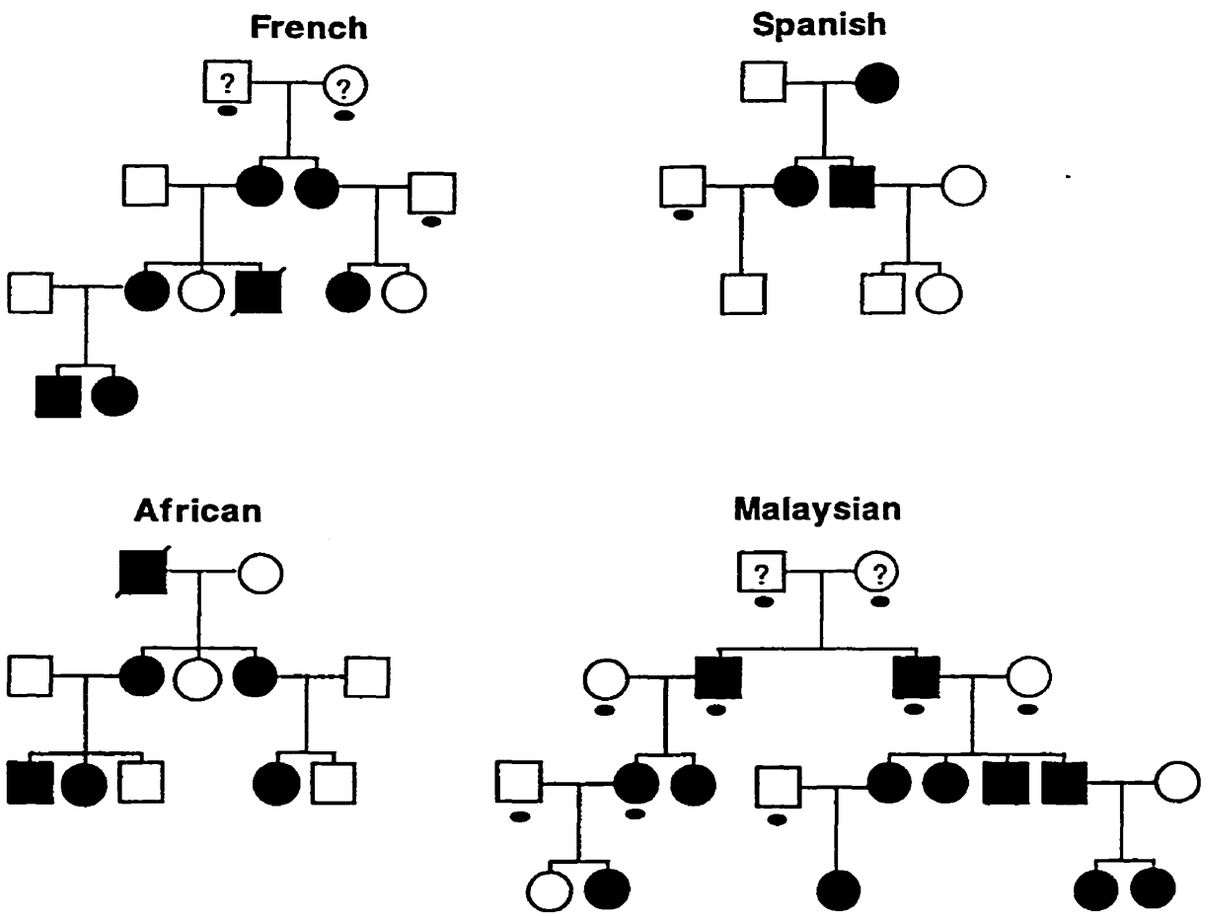

Figure 5. 4. Pedigrees of the French, Spanish, African and Malaysian families used in this study. All live members have DNA available for analysis except those indicated by black ovals.

This order was examined by the CRI-MAP program developed by Phil Green (Lander and Green, 1987) using combined data from 14 CEPH reference families and 20 HED families. We first ran the Chrompic option to find all recombinants in the region; apparent double crossovers were considered as candidate data errors and the genotyping was repeated accordingly. We next used the Fixed option to find the associated maximum likelihood recombination fractions and map distances between the ordered loci and the Flips2 option to test permutations of adjacent loci within this order. The order of these markers and intermarker genetic distances in Kosambi cM are as follows: D13S1826-0.2-D13S1828-0.5-D13S1827-0-D13S1835-0.8- D13S141-0-D13S175-1.1-D13S1830-0-D13S1831-0-D13S1832-2.5-D13S143. The relative $\log_{10}$-likelihoods of 2.13, 4.29, 5.27, 5.10 and 7.92 were obtained with permutations of loci D13S1826-D13S1828, D13S1828-D13S1827, D13S1835-D13S141, D13S175-D13S1830 and D13S1832-D13S143 respectively. This order was compatible with data from haplotype analysis. For markers not separated by crossover events in our sample, a fixed intermarker recombination fraction of 0.01 cM was assumed for multipoint analysis.

Linkage Analysis

Two-point linkage analysis was performed using the MLINK and ILINK programs of the LINKAGE 5.1 package (Lathrop *et al.*, 1984; Terwilliger and Ott, 1994). Multi-point analysis was done using the VITESSE program developed by O'Connell and Weeks (1995). The mode of inheritance was considered to be autosomal dominant with complete penetrance and an estimated gene frequency of 0.00001. Recombination

fractions were assumed to be equal in females and males. Allele frequencies for all markers were obtained from the genotypes of 50 CEPH control individuals.

Linkage Disequilibrium Analysis

French Canadian Sample

Four French Canadian families are related 4 generations back and hence were considered as one family for linkage disequilibrium analysis. Consequently, only 20 French Canadian families were available for this analysis. One affected individual carrying the full disease haplotype from each of these families was selected for analysis; except in those families where a recombination event took place, then two or more affected haplotypes were included accordingly. This gave rise to a total of 26 affected HED chromosomes. The normal population allele frequency was obtained from the genotypes of 29 unaffected spouses participating in this study.

Allelic association

Non-random allelic association between the markers and the HED gene in the French-Canadian affected chromosomes was confirmed using Fisher's exact test on a 2×2 table. The allele with the highest frequency in the affected chromosomes was assumed as the associated allele and all other alleles were combined to form the second allele. The strength of association was determined using the formula: $P_{excess} = (P_{affected} - P_{normal})/(1 - P_{normal})$ where P_{normal} and $P_{affected}$ represent the allele frequencies in the general population and disease-carrying chromosomes respectively (Lehesjoki *et al.*, 1993).

Multipoint Linkage Disequilibrium Analysis

A multipoint linkage disequilibrium analysis was performed using the DISMULT version 2.1 routine of the likelihood method developed by J. D. Terwilliger (18). Ten markers were analyzed simultaneously according to the map order and distances mentioned above. The minimum and maximum values for the decay parameter representing the number of generations since the introduction of the disease chromosome in the population was set to be from 5 to 20. The number of test points for theta in the inter-marker intervals was set to 10, and the allele frequency for the disease gene in was set at 1 in 10000 according to the estimated number of cases observed in the population.

ACKNOWLEDGMENTS

We would like to thank the families that made this study possible. This work was supported by grants from the Canadian Dermatology Foundation (G.A.R), the National Foundation for Ectodermal Dysplasias (G.A.R and J.Z.) and NIDR DE11311 (J.Z.). G.A.R is supported by the Medical Research Council of Canada.

REFERENCES

Ando, Y., Tanaka, T., Horiguchi, Y., Ikai, K., and Tomono, H. (1988). Hidrotic ectodermal dysplasia. *Dermatologica* **176:** 205-211.

Boleda, M. (1984) Les migrations au Canada sous le régime français (1608-1760). *Cah. Québécois Demogr.* **13:** 23-39.

Bouchard, G., and De Braekeleer. (1991). Mouvements migratoires, effets fondateurs et homogénéisation génétique. In *Histoire D'un Génome*. Presses de l'Université de Québec, Québec, pp.283-321.

Brais, B., Bouchard, J. P., Jomphe, M., Desjardins, B., Dubé, M. -P., Gosselin, F., Xie, Y. -G., Tomé, F. M. S., Fardeau, M., Brunet, G., Samson, F., Mathieu, J., Prévost, C., Duranceau, A., Codère, F., Pericak-Vance, M., Bulman, D.E., Tapscott, S., Feasby, T., Munsat, T. L., Pilarski, R., Morgan, K., and Rouleau, G.A. (1998). When genetics and history converge: the fine-mapping and North American introduction and diffusion of the French Canadian oculopharyngeal muscular dystrophy mutation. *Am. J. Hum. Genet.* **In press.**

Casaubon, L. K., Melanson, M., Lopes-Cendes, I., Marineau, C., Andermann, E., Andermann, F., Weissenbach, J., Prévost, C., Bouchard, J. -P., Mathieu, J., and Rouleau, G.A. (1996). The gene responsible for a severe form of peripheral neuropathy and agenesis of the corpus callosum maps to chromosome 15q. *Am. J. Hum. Genet.* **58:** 28-34.

Christianson, A.L., and Fourie, S. (1996). Family with autosomal dominant hidrotic ectodermal dysplasia: a previously unrecognized syndrome? *Am. J. Med. Genet.* **63:** 549-553.

Clouston, H. R. (1929). Hereditary ectodermal dystrophy. *Can. Med. Ass. J.* **21:** 18-31.

Clouston, H.R. (1939). The major forms of hereditary ectodermal dysplasia. *Can. Med. Ass. J.* **40:** 1-7.

De La Chapelle, A., and Wright, F.A. (1998) .Linkage disequilibrium mapping in isolated populations: the example of Finland revisited. *Proc. Nat. Acad. Sci. USA* **95:** 12416-12423.

Der Kaloustian, V. M., and Kurban, A.K. (1979). In *Genetic Diseases of the Skin.* Springer-Verlag, pp. 109-114.

Desjardins, B. (1990). Homogénéité ethnique de la population québécoise sous le régime français. *Cah. Québécois Demogr.* **19:** 63-76.

Devlin, B., and Risch, N. (1995). A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* **29:** 311-322.

Devlin, B., Risch, N., and Roeder, K. (1996). Disequilibrium mapping: composite likelihood for pairwise disequilibrium. *Genomics* **36:** 1-16.

Escobar, V., Goldblatt, L. I., Bixler, D., and Weaver, D. (1983). Clouston sydrome: an ultrastructural study. *Clin. Genet.* **24:** 140-146.

Giraud, F., Mattei, J. F., Rolland, M., Ghiglione, C., Pommier de Santi, P., and Sudan, N. (1977). La dysplasie ectodermique de type Clouston. *Arch. Franç. Péd.* **34:** 982-993.

Gold, R. J. M., and Scriver, C. R. (1972). Properties of hair keratin in an autosomal dominant form of ectodermal dysplasia. *Am. J. Hum. Genet.* **24:** 549-561.

Hayflick. S. J., Taylor, T., McKinnon, W., Guttmacher, A. E., Litt, M., and Zonana, J. (1996). Clouston syndrome (hidrotic ectodermal dysplasia) is not linked to keratin gene clusters on chromosomes 12 and 17. *J. Invest. Dermatol.* **107:** 11-14.

Heyer, E. (1995). Mitochondrial and nuclear genetic contribution of female founders to a contemporary population in Northeast Québec. *Am. J. Hum. Genet.* **56:** 1450-1455.

Hill, W.G., and Weir, B. S. (1994). Maximum-likelihood estimation of gene location by linkage disequilibrium. *Am. J. Hum. Genet.* **54:** 705-714.

Jorde, L. B. (1995). Linkage disequilibrium as a gene-mapping tool. *Am. J. Hum. Genet.* **56:** 11-14

Kibar, Z., Der Kaloustian, V. M., Brais, B., Hani, V., Fraser, F. C., and Rouleau, G.A. (1996). The gene responsible for Clouston hidrotic ectodermal dysplasia maps to the pericentromeric region of chromosome 13q. *Hum. Mol. Genet.* **5:** 543-547.

Kibar, Z., Lafrenière, R. G., Chakravarti, A., Wang, J. -C., Chevrette, M., Der Kaloustian, V. M., and Rouleau, G.A. (1999). A radiation hybrid map of 48 loci including the Clouston hidrotic ectodermal dysplasia (HED) locus in the pericentromeric region of chromosome 13q. *Genomics* **56:** 127-130.

Labuda, M., Labuda, D., Korab-Laskowska, M., Cole, D. E. C., Zietkiewicz, E., Weissenbach, J., Popowska, E., Pronicka, E., Root, A. W., and Glorieux, F. H. (1996). Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. *Am. J. Hum. Genet.* **59:** 633-643.

Lander, E., and Green, P. (1987). Construction of multilocus genetic linkage maps in humans. *Proc. Natl. Acad. Sci. USA* **84:** 2363-2367.

Lathrop, G. M., Lalouel, J. M., Julier, C., and Ott, J. (1984). Strategies for multilocus linkage analysis in humans. *Proc. Nat. Acad. Sci. USA* **81:** 3443-3445.

Lehesjoki, A. -E., Koskiniemi, M., Norio, R., Tirrito, S., Sistonen, P., Lander, E., and De La Chapelle, E. (1993). Localization of the *EPM1* gene for progressive myoclonus

epilepsy on chromosome 21: linkage disequilibrium allows high resolution mapping. *Hum. Mol. Genet.* **2:** 1229-1234.

Mcnaughton, P. Z., and Pierson, L. (1976).. Hidrotic ectodermal dysplasia in a black mother and daughter. *Arch. Dermatol.* **112:** 1448-1450.

O'Connell, J. R., and Weeks, D. E. (1995). The VITESSE algorithm for rapid exact multilocus linkage analysis via genotype set-recoding and fuzzy inheritance. *Nature Genet.* **11:** 402-408.

Patel, R. A, Bixler, D., and Norins A.L.(1991). Clouston syndrome: a rare autosomal dominant trait with palmoplantar hyperkeratosis and alopecia. *J. Craniofac. Gen. Dev. Biol.* **11:** 176-179.

Pierard, G. E., Van Neste, D., and Letot, B. (1979). Hidrotic ectodermal dysplasia. *Dermatologica* **158:** 168-174.

Radhakrishma, U., Blouin, J. -L., Mehenni, H., Mehta, T. Y., Sheth, F. J., Sheth, J. J., Solanki, J. V., and Antonarakis, S. E. (1997). The gene for autosomal dominant hidrotic ectodermal dysplasia (Clouston syndrome) in a large Indian family maps to the 13q11-q12.1 pericentromeric region. *Am. J. Med. Genet.* **71:** 80-86.

Rajagopalan, K., and Hai Tai,C. (1977). Hidrotic ectodermal dysplasia. *Arch. Dermatol.* **113:** 481-485.

Scott, D. A., Carmi, R., Elbedour, K., Duyk, G. M., Stone, E. M., and Sheffield, V. C. (1995). Nonsyndromic autosomal recessive deafness is linked to the DFNB1 locus in a large inbred Bedouin family from Israel. *Am. J. Hum. Genet.* **57:** 965-968.

Stevens, H. P., Choon, S. E., Hennies, H.-C., and Kelsell, D. P. Genetic Homogeneity in Clouston's hidrotic ectodermal dysplasia. *Br. J. Dermatol.* **In press.**

Taylor, T. D., Hayflick, S. J., McKinnon, W., Guttmacher, A. E., Hovnanian, A., Litt, M., and Zonana, J. (1998). Confirmation of linkage of Clouston syndrome (hidrotic ectodermal dysplasia) to 13q11-q12.1 with evidence for multiple independent mutations. *J. Invest. Dermatol.* **111:** 83-85.

Terwilliger, J. D. (1995). A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am. J. Hum. Genet.* **56:** 777-787.

Terwilliger, J. D., and Ott, J. (1994). In *Handbook of Human Genetic Linkage*, the John Hopkins University Press, London, pp. 22-32.

Warburton, D., Shaw, S. H., Matise, T. C., Kalachikov, S., and Fischer, S. (1996) Report of the third international workshop on human chromosome 13 mapping 1995. *Cytogenet. Cell Genet.* **75:** 85-110.

Worobec-Victor, S. M., Bene-Bain, M. A., Shanker, D. B., and Solomon, L. M. (1988). In Schachner, L. A. and Hansen, R. C. (eds.), *Pediatric Dermatology.* Churchill Livingstone, New York, Vol.I, p. 328.

Xiong, M., and Guo, S. W. (1997). Fine-scale genetic mapping based on linkage disequilibrium: theory and applications. *Am. J. Hum. Genet.* **60:** 1513-1531.

CHAPTER 6

Conclusion

6.1. SUMMARY

Clouston Hidrotic etctodermal dysplasia (HED) is a form of genodermatosis transmitted as an autosomal dominant condition with complete penetrance and variable expressivity (Williams and Fraser, 1967). The main clinical features associated with HED include nail dysplasia, defects of the hair and palmoplantar hyperkeratosis (Der Kaloustian and Kurban, 1979). While this disease is rare worldwide, it is particularly common in the French Canadian population with an estimated gene frequency of 1 in 10000 (Clouston, 1929; Kibar *et al.*, submitted). The molecular defect underlying HED has not been identified yet and this study establishes the basic steps towards achieving this goal.

6.1.1. Linkage mapping of the HED gene

Pathologic studies of the defective hair and nails in HED suggest a molecular abnormality of keratins or of keratin-associated proteins (KAP) (Escobar *et al.*, 1983; Giraud *et al.*, 1977; Gold and Scriver, 1972). This abnormal keratinization could be the primary genetic defect caused by mutations in keratin or in KAP genes or secondary to inappropriate developmental signals leading to faulty hair and nail proliferation and/or differentiation (Holbrook, 1988). We tested eight French Canadian families segregating HED for linkage to polymorphic markers spanning regions on chromosomes 1, 11, 12 and 17 that contain the known keratin and KAP genes and we excluded these areas. Another study by Hayflick *et al.* (1996) confirmed exclusion of HED to candidate regions on chromosomes 1, 12 and 17 and extended this exclusion to two other regions: 18q21 that contains three desmoglein genes (Simrak *et al.*, 1995) and to which striate palmoplantar

keratoderma was mapped (Hennies et al., 1995), and 20q11.2 that contains the gene coding for transglutaminase 3 (Wang *et al.*, 1994). Even though both studies excluded linkage of HED to the known keratin and KAP genes, HED could still be caused by a mutation in an unlocalized keratin or KAP gene. Recent data indicate that hair keratins constitute a complex family of proteins, eleven members of which have been characterized (seven type-I and four type-II) (Winter et al., 1997). The type II subfamily is still incomplete and even though it is expected that the missing members are clustered with the other functionally and sequentially related type-II genes on chromosome 12q13, they could have divergent sequences and map to other regions of the human genome. The hair-specific KAPs constitute a large family of at least 50 genes best studied in the sheep genome (Powell *et al.*, 1991) and only four of these have been characterized in humans (Emonet *et al.*, 1997, Mackinnnon *et al.*, 1990)

In the absence of other biochemical defects specific to HED, a positional cloning strategy was initiated to identify the defective gene. The first step, which is mapping the HED gene to a chromosomal region, was achieved by linkage analysis in eight French Canadian families. A genome scan using well-mapped Genethon polymorphic markers was initiated, and after excluding an additional 10% of the genome, linkage was found to the pericentromeric marker D13S175 with a maximum combined two-point lod score of 8.12 at $\theta = 0$. Lod scores above 3.00 were obtained with six other chromosome 13q pericentromeric markers confirming linkage of HED to this region. Haplotype analysis using the markers cent-D13S175-D13S141-D13S143-D13S115-D13S1236-D13S292-D13S283-tel defined the candidate region to extend from the centromere to D13S143. Two other studies confirmed linkage of HED to this region in families of Indian

(Radhakrishma *et al.*, 1997) and Scottish-Irish (Taylor *et al.*, 1998) origins, thereby suggesting absence of non-allelic genetic heterogeneity in HED. Later, we have extended this finding of non-allelic genetic homogeneity in HED using four families of African, French, Spanish and Malaysian origins. Affected chromosomes in this multi-ethnic group of families carry different haplotypes and therefore most likely carry different mutations of the HED gene. The existence of different ancestral HED mutations is consistent with findings in other diseases.

6.1.2. Exclusion of genes co-localized to the HED gene for involvement in HED

The HED candidate region established in the initial linkage study contains five known genes: connexin 26 (*Cx26*), connexin 46 (*Cx46*) (Mignon *et al.*, 1996), fibroblast growth factor 9 (*FGF-9*) (Mattei *et al.*, 1995), zinc finger 198 (*ZNF198*) (Reiter *et al.*, 1998) and α-tubulin *TUBA2* (Guilford *et al.*, 1995). These genes were considered as candidates for HED and hence were subjected to mutation screening in HED patients. The human cDNA sequence was available for all of them except Cx46. Using combined efforts from our group and the group of Dr. Alain Shiels at Washington University (St-Louis), the human *Cx46* coding sequence was cloned by homology to rodent *Cx46* and human *Cx50* cDNA sequences. The open reading frame (ORF) of each of the five genes and the exon-intron junctions of *TUBA2* were screened by PCR-SSCP analysis for the presence of basepair alterations specific to HED patients and none was found. The most likely exclusion of *Cx26*, *Cx46* and *FGF-9* for involvement in HED by PCR-SSCP analysis was later supported by the findings of pathogenic mutations in *Cx26* and *Cx46* in two other diseases, hereditary non-syndromic sensorineural deafness (Kelsell *et al.*, 1997)

and autosomal dominant "zonular pulverulent" cataract (Mackay *et al.*, 1999) respectively, and by later exclusion of *FGF-9* from the HED candidate region by recombinational analysis. Our PCR-SSCP results with the two other genes, *ZNF198* and *TUBA2*, suggest that they are most likely not defective in HED. If further genetic analysis does not exclude these two genes from the HED candidate region, efforts will be focused on detailed analysis of their sequences for mutations in HED.

In addition to these genes, the HED candidate region contained 18 expressed sequence tags (ESTs) (shown in figure 4.2.). Using the Blast family of tools (Blastn, Blastp, Blastx and tBalstn), these ESTs were screened to identify homologous sequences and no homology relevant to skin biology was found. Consequently, they were not further pursued as candidates for HED. The further refined HED candidate region flanked by D13S1828 and D13S1830 contains only three of these ESTs, A001U40, D13S183E, and WI-11739.

6.1.3. Construction of a radiation hybrid map including the HED locus

The HED candidate region established in the initial mapping report spanned at least 15 cM and is too large for positional cloning efforts. Finer localization of the HED gene requires the collection of additional families and the isolation of new polymorphic markers to increase the number of informative meioses and the likelihood of finding key recombinants. A long-range map is essential to serve as a basis for isolation of these markers and of new genes from the region and to provide order and distance estimates for these loci. While genetic mapping is useful in constructing long-range maps of the human genome, it is linkage-based and hence can only incorporate polymorphic markers (Matise

et al., 1994). The cloning system of yeast artificial chromosomes (YACs) can accommodate DNA inserts of large sizes (up to 1 Mb or more) and has been used along with pulse field gel electrophoresis to assemble long-range contigs for physical mapping purposes. However, this system has the disadvantages of clone instability, chimerism and chromosomal rearrangements and deletions (Schlessinger, 1990). In addition, the construction of YAC contigs in our candidate region will be complicated by the high sequence homology reported between the pericentromeric regions of chromosomes 13 and 21 (Van camp *et al.*, 1992). In fact, Guilford *et al.* (1995) have constructed a YAC contig in the pericentromeric region of chromosome 13q including the loci for neurosensory nonsyndromic deafness and limb-girdle muscular dystrophy type 2C and have anticipated the difficulty in completing their YAC contig up to the centromere because of this homology. Another long-range physical mapping tool is radiation hybrid (RH) mapping that can achieve a level of resolution between genetic mapping on one hand and YAC contig building on the other hand. This method can incorporate loci derived from any type of DNA sequence, polymorphic or monomorphic, thereby integrating data from available physical and genetic maps (Cox *et al.*, 1990). In addition, by using clones derived from a somatic cell hybrid containing only human chromosome 13 in a rodent background, the problems associated with sequence homology to chromosome 21 can be avoided. The compromising factors of YAC cloning and the simplicity and advantages of RH mapping led us to choose the latter to build a physical map of the HED candidate region.

RH data analysis was initially started with 54 loci derived from the available maps of chromosome 13q. Two-point analysis identified one linkage group of 49 loci including

the centromeric marker D13Z1 and the distal flanking marker for the HED candidate region D13S143. Multi-point analysis of this group was first carried out under an equal retention model and generated a 1000:1 framework map of 30 loci. Examination of the RH data of these loci revealed a decrease in retention frequency as a function of distance from the marker D13Z1. This "centromeric effect" has been reported in other studies and was expected to be obtained with our data. Consequently, multi-point RH analysis was rerun under a centromeric retention model and a 1000:1 framework map of 29 loci was generated, that is ~2000 times more likely than the next order. The resulting map was dependent on excluding D13Z1 from the analysis. In fact, when this locus was included, only 3 loci could be ordered with 1000:1 level of support. D13Z1 represents highly repetitive alpha satellite DNA and hence it does not detect one locus (Greig *et al.*, 1993); this fact along with its high retention frequency (65.6%) could have had complicated the analysis. Importantly, when the 29 loci forming the framework map under a centromeric model were reanalyzed under an equal retention model, the same best order was obtained but with different total map length and intermarker distance estimates. The centromeric model fit the data significantly better than the equal retention model and consequently we considered the framework map obtained under the first model as the correct one. The other loci that could not be ordered with this level of support were placed in a range of adjacent intervals. This map spanned 347 cR_{9000} with an average resolution of 17.3 cR_{9000} (~865 Kb-1.3 Mb). It integrates 3 genes (*TUBA2*, *GJβ2* and *FGF-9*) and 18 ESTs with 27 other markers including 19 polymorphic markers.

Comparison of our RH map to the detailed YAC contig and EST map constructed by Guilford *et al.* (1995) in the region revealed a reversed interval of six loci in the YAC

contig. The YAC contig order was later confirmed by Reiter *et al.* (1998) as part of a positional cloning strategy to isolate the gene fused to the fibroblast growth factor receptor-1 in the t(8;13)(p11;q12). The likelihoods for the two orders were compared with RHMAP and the RH order was found to be the best locus order with a log likelihood difference of 17 from the YAC contig order. Fiber-FISH and FISH on interphase nuclei analyses with five PACs isolated from the region supported our order. In addition, genetic analysis of 10 polymorphic markers isolated from the region (eight of which were isolated in our laboratory) with the program CRI-MAP (Lander and Green, 1987) in a set of 14 CEPH control families and 20 HED families further supported our order with log10likelihood difference of ~10. This inconsistency in order could be caused by YAC rearrangements and/or by the extensive homology reported between chromosomes 13 and 21 pericentric regions especially that PCR results in both studies were not confirmed by Southern blot analysis. A high level of sequence homology between these two interchromosomal regions was encountered in our study where 4 out of 8 PACs isolated with our most centromeric marker D13S1118 were of chromosome 21 origin.

6.1.4. Isolation of new polymorphic markers in the HED candidate region

Based on a preliminary RH map, 13 markers from the HED candidate region were selected to generate additional $(CA)_n$ polymorphic markers in this region. They were screened against a human P1 genomic library and 57 PACs were isolated. Cross-hybridization with sequences other than those on chromosome 13q was encountered and only 22 clones were found to be of chromosome 13 origin. Small insert libraries generated for these 22 PACs, 4 chromosome 21 PACs (isolated with chromosome 13q

markers) and 4 YACs that partially cover the HED candidate region were screened with a (polydA-dC).(polydG-dT) probe and eight new polymorphic $(CA)_n$ markers were developed. Two markers had PIC values > 0.80 and three markers > 0.70, which makes them highly useful in the genetic mapping of HED as well as other diseases mapped to the region.

6.1 5. HED: demonstration of a founder effect in the French Canadian population and fine genetic mapping using linkage and linkage disequilibrium analyses

To refine the localization of the HED gene, seven of the newly isolated markers along with three others were used for genetic analysis of a multi-ethnic group of 29 HED families. The order of the markers used and intermarker genetic distances in Kosambi cM were as follows: cent- D13S1826-0.2-D13S1828-0.5-D13S1827-0-D13S1835-0.8-D13S141-0-D13S175-1.1-D13S1830-0-D13S1831-0-D13S1832-2.5-D13S143-tel. The order of these markers was obtained from the RH map we constructed in the region and was analyzed with the CRIMAP program using genotyping data from 14 CEPH control families and 20 HED families. The minimum co-segregating region was defined by linkage and linkage disequilibrium (LD) analysis.

Haplotype analysis detected nine recombinants with the two most critical recombinants placing the HED gene in a 2.4 cM region flanked by D13S1828 proximally and D13S1830 distally. A multipoint linkage analysis was also carried out to use information in all 29 HED families and a maximum lod score of 52.13 was obtained at D13S1835. The one-lod-unit support interval spanned 1.93 cM, 0.05 cM telomeric to D13S1828 and 0.43 cM centromeric to D13S1830.

Given the particular history of the French Canadian population, all HED-affected chromosomes in the present population are most likely descendants from one or few founder chromosomes that carried the HED mutation. Consequently, all affected chromosomes would share the same haplotype and show a strong allelic association with markers in the vicinity of the HED mutation. This "founder effect" hypothesis was confirmed with Fisher exact test and a significant allelic association was found between the 10 markers and the affected HED chromosomes ($P<0.001$). The strongest allelic association was detected between HED and alleles at markers D13S1828, D13S1827, D13S1835, D13S141 and D13S175 ($P_{excess} = 1$). The affected chromosomes were next examined for the presence of an ancestral haplotype that most likely carried the founding mutation and of historical recombinants that would help us map the HED gene. A major haplotype that most likely represents the founder haplotype including all 10 associated alleles was present on 65% of affected chromosomes. Nine historical recombinants were detected placing the HED gene in a 1.3 – 2.6 cM region including the markers D13S1828, D13S1827, D13S1835, D13S141 and D13S175. Consequently, a multipoint LD method developed by J. D. Terwilliger (1995) was applied to estimate the most likely position of the HED gene. This method uses all alleles at each marker locus without assumption of the associated allele. The pairwise information is combined by multiplication for larger number of loci in a multipoint likelihood analysis. Our LD multipoint analysis placed the HED gene most likely at 0.08 cM telomeric to D13S1835 and 0.72 cM centromeric from D13S141 with a support interval spanning 3.27 cM.

The multipoint likelihood method (Terwilliger, 1995) used in our study does not make assumptions about the evolutionary history of the population. According to Devlin

et al. (1996), this method overestimates the strength of the data by not taking into account the evolutionary sampling process. In contrast, Kaplan *et al.* (1995) used a Poisson branching process to simulate disease chromosome populations. Samples are drawn from the simulated populations for likelihood calculations of the recombination fraction between marker and disease gene. This model does not generalize when multiple marker loci are available; it can accommodate multipoint analysis of all alleles only for two markers. Another LD mapping method commonly used to fine-map disease genes in founder populations is the Luria- Delbrück model originally developed to estimate mutation rates in bacteria. Hästbacka et al. (1992) applied this model on human populations to estimate recombination rates. This model assumes exponential growth of the population and directly uses the proportion of non-recombinants within the population to estimate the genetic distance between a disease locus and a closely linked marker. Both disease and marker loci are assumed to be biallelic. Estimation of the confidence limits for the recombination fraction is based on a crude calculation of the likely mean and standard deviation of the proportion of recombinant chromosomes in the present population.

The performance of the three methods described above was examined in a dominant disease particularly common in the French Canadian population called oculopharyngeal muscular dystrophy (M.-P. Dubé, personal communication). The OPMD gene was localized by simple LD mapping to a 0.4 cM region flanked by D14S990 proximally and D14S1457 (Brais *et al.*, in press). The defective gene was later identified and found to map at ~210 Kb telomeric to D14S990 (Brais *et al.*, 1998). While both Luria- Delbrück and multipoint likelihood analyses performed fairly well with the data,

the multipoint Poisson branching process likelihood method was the model that best estimated the position of the OPMD gene based on LD data. This model positioned the gene at 0.26 cM (~260 kb) telomeric to D14S990 with a support interval of 0.21 cM within the D14S990-D14S1457 interval. Consequently, this kind of analysis was redone with HED and the gene was most likely positioned at 0.05 cM telomeric to D13S1828 with a support interval of 4.0 cM extending from 2.06 cM centromeric to D13S1826 to 0.65 cM centromeric to D13S1830. Figure 6.1. summarizes the results obtained with linkage and linkage disequilibrium analyses. As shown in this figure, the HED gene most likely resides in the proximal part of the interval D13S1828-D13S1830 that is defined by recombination mapping.

6.2. Future prospects

The HED candidate region spans 2.4 cM that is still large for positional cloning efforts. Further narrowing down this region by recombination mapping requires the collection of additional HED families and isolation of new polymorphic markers. Results obtained with multipoint linkage and LD analyses will direct the physical mapping strategies by aiming at constructing a physical contig around the markers D13S1828, D13S1827 and D13S1835. The contig will help estimate the physical size of the candidate region more precisely, order the markers more accurately and will serve as a genomic basis for isolation of additional polymorphic markers and transcription units from the region. Once the genetic and physical distances have been further defined, classic positional cloning methods can be applied to isolate the defective gene.

The process of positionally cloning the HED gene from a defined physical region can be quite laborious depending on the size of this region and on its gene-density. Even

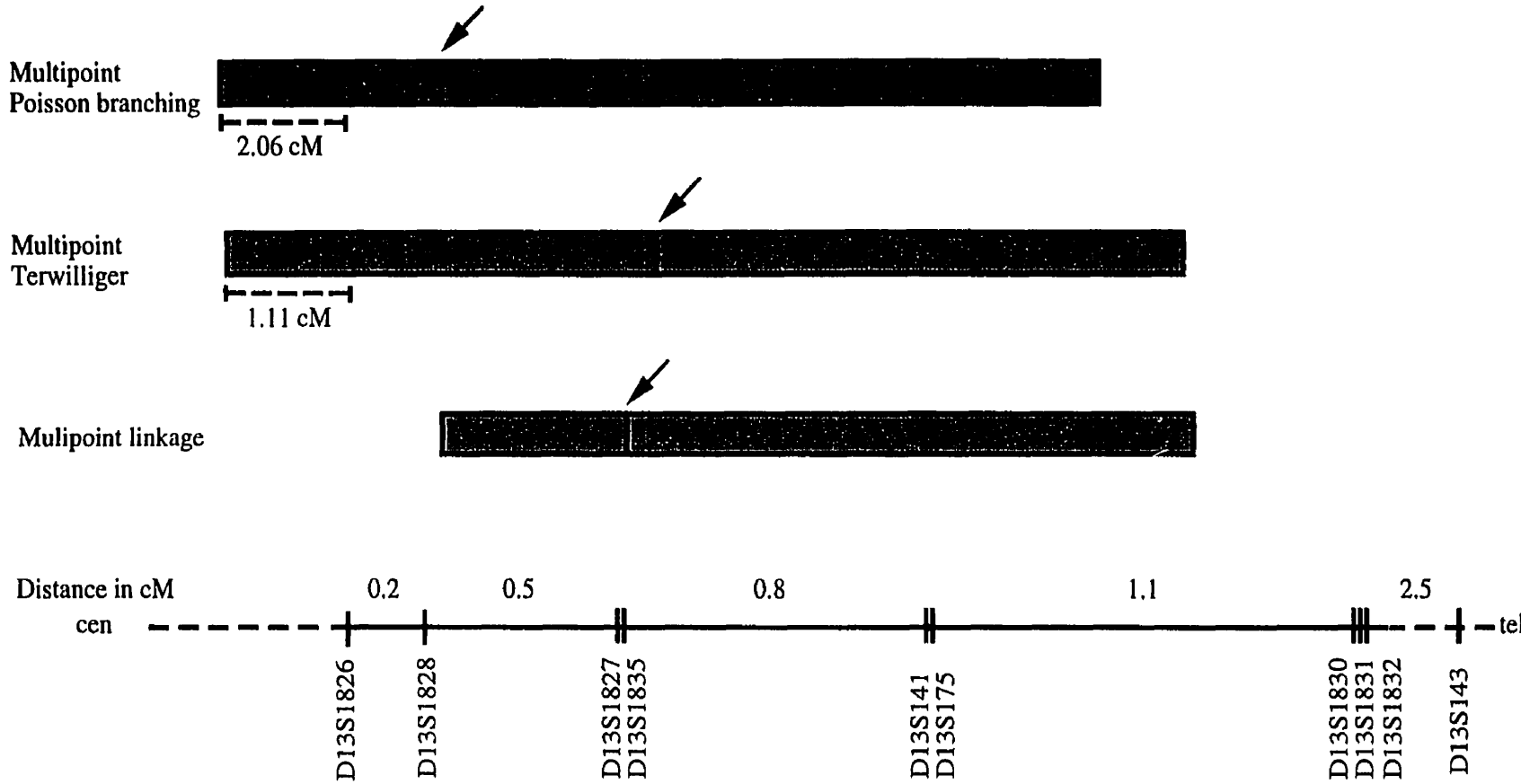

Figure 6. 1. Summary of the results obtained with linkage and linkage disequilibrium analyses. Arrows and shaded boxes respectively indicate the most likely positions of the HED gene and support intervals obtained by these analyses.

though chromosome 13 has been shown to be one of the least GC rich and presumably least gene-rich chromosomes (Antonarakis, 1994; Craig and Bickmore, 1994), the HED candidate region could represent an intrachromosomal exception to this general observation and be highly dense with genes. An example of such intrachromosomal variation exists on chromosome 21 where band 21q22.3 is very rich in CpG islands compared to the rest of the chromosome (Antonarakis, 1994). Identification of the HED gene could be accelerated by the following factors:

1) positional candidate gene approach and the human genome project

The positional candidate gene approach has been successfully used in cloning a number of human disease genes (Collins, 1995). For example, a group of bone malformation syndromes, Pfeiffer syndrome, Crouzon syndrome and achondroplasia, were first mapped by linkage analysis to chromosomes 8p, 10q and 4p, respectively. Three fibroblast growth factor receptor (FGFR) genes, FGFR1, FGFR2, and FGFR3, were physically mapped to these regions. Subsequent searches and findings of mutations in these genes in affected individuals have demonstrated their involvement in these diseases (Wilkie *et al.*, 1995).

The rate at which cDNA sequences and expressed sequence tags (ESTs) are being generated and deposited in the databases presently is very high. In fact, the Human Genome Project, designed to produce a complete nucleotide sequence map of the human genome, has major focuses on producing a dense human transcript map with mapping and sequence information on thousands of ESTs. Achieving this goal will make the positional candidate approach dominate the other approaches in identifying disease genes (Collins, 1995; Lander, 1996).

2) human-mouse homology

Conservation of genomic organization in mammalian species has long been established and is studied in details bewteen human and mouse. Knowledge of this homology can be extremely useful in mapping and identifying genes in both species. Mapping or cloning a gene in one species will help predict the localization of or isolate the orthologous gene in the other species. The HED candidate region located at 13q11-q12 contains two genes, connexin 26 and connexin 46, which map to mouse chromosome 14D1-E1 indicating conservation of synteny in these subchromosomal regions (Mignon *et al.*, 1996). The rapid progress in mapping and generation of ESTs in human and mouse will expand information about this synteny and could help identify genes located in these regions.

The presence of an animal model of a human disease where the causative mutations have been defined might also help identify the human gene. The usefulness of this approach is limited by species differences in some of the biological processes including differences in some biochemical and developmental pathways. An ectodermal malformation in mice bearing the dominantly inherited Naked trait has been described to have a phenotype similar to HED. Heterozygous mice (Nn) show breakage of the hair close to the root leading to patchy depilation. Homozygous mice (NN) have almost no hair or nail and usually die in utero or within 10 days of birth. Tenenhouse et al (1976) demonstrated an abnormal protein composition of hair in Nn mice. They found a reduction in the amounts of glycine and tyrosine that are probably indirect effects of the *N* mutation. The N gene was localized to the distal end of mouse chromosome 15 that is

equivalent to 12q in human (Saffer *et al.*, 1990) and hence this gene is most likely not the orthologue of the human HED gene. Therefore, no mouse model of HED currently exists.

3) Homology to a gene involved in a similar human disease phenotype

An identified disease gene might belong to a multigene family where mutations in other members could cause diseases with similar phenotypes. Two such genes responsible for X-linked anhidrotic ectodermal dysplasia (EDA) and recessive alopecia universalis (AU) could be relevant to the isolation of the HED gene by using this functional homology.

EDA is a form of ectodermal dysplasia characterized by sparse hair, abnormal or missing teeth and inability to sweat due to lack of sweat glands. The skin does not show any specific histological changes but there is an overall decrease in the number, size and maturation of eccrine sweat glands, sebaceous glands and hair follicles. The gene has been positionally cloned and found to encode a predicted 391-amino-acid protein (Kere *et al.*, 1996; Monreal *et al.*, 1998). The gene expression starts as early as Week 8 of gestation in some tissues and continues through fetal life with a staged timing pattern to some adult tissues mainly heart, pancreas, prostate and uterus. It is also expressed in keratinocytes, hair follicles (outer root sheath cells) and sweat glands. The predicted protein includes a transmembrane helix-forming collagen-like domain and is likely to be a cell-surface protein. It is believed to belong to a novel class with a role in epithelial-mesenchymal signaling or in cell adhesion or cell migration (Monreal *et al.*, 1998; Montonen *et al.*, 1998).

AU is a form of hereditary hair loss which results in the complete absence of scalp and body hair. The gene was mapped to chromosome 8p12 by homozygosity mapping

and then cloned by homology to the *hairless* gene in mouse. The human hairless encodes a putative protein of 1189 amino acids with restricted expression in the brain and skin. The protein contains a single zinc finger with a novel and conserved six-cysteine motif and thought to act as a transcription regulator at the cellular transition from the natal to the first adult hair cycle (Ahmad *et al.*, 1998a; 1998b).

Both EDA and hairless genes appear to belong to novel families with regulatory roles in the development of the epidermal appendages. HED is also believed to be the result of a developmental error in the epithelial-mesenchymal inductive process (Holbrook, 1988) and hence it might be functionally related to either of these two genes. This hypothesis could be tested by trying to isolate the HED gene using either reduced stringency hybridization with probes consisting of full length EDA or hairless cDNA or PCR amplification using degenerate primers developed to the most conserved domains among these genes.

6.3. Implications

This study constitutes the essential steps in the molecular characterization of the defect causing HED. Identification of the responsible gene and characterization of its protein product might lead to the development of animal models which could elucidate the pathophysiology of HED and provide the basis for better therapeutic strategies. In addition, studies of the HED gene might provide insights into some of the molecular mechanisms underlying the normal developmental and physiological processes of the hair and nail appendages. Skin appendages such as hair and nail are formed through a multistage process of complex interactions between the epithelium and mesenchyme, that

involves various morphogens with major regulatory roles in development (Holbrook *et al.*, 1993; Stenn *et al.*, 1996). In addition to the complexity of these signaling pathways, the hair cortex and cuticle cells express an estimated number of 50-100 proteins belonging to at least 10 distinct families (Powell *et al.*, 1991) and the nail structures develop with more complex patterns of keratin expression (Moll *et al.*, 1988). The molecular events implicated in these processes are still poorly understood and characterization of genes defective in ectodermal dysplasias might help unravel some of them.

This study also presents data that are useful for the mapping and isolation of other disease loci in the pericentromeric region of chromosome 13q. It describes a detailed radiation hybrid map of 48 loci in the region including eight new polymorphic markers and provides supporting evidence for the order of six of these loci that was found inconsistent with a YAC contig order. Establishment of the correct order by our RH map will help better understand the nature of the chromosomal abnormality t(8;13)(p11;q12) commonly associated with the 8p11 myeloproliferative disorder. This translocation is believed to be a complex rearrangement involving an inversion of the region on chromosome 13q11-q12 based on the YAC contig order (Reiter *et al.*, 1998).

Finally, this study represents the first example of a founder disease in the French Canadian population in the southwestern part of the province of Québec. Results of the linkage disequilibrium analysis illustrate the suitability of the French Canadian population for the fine mapping of founder disease loci. A major asset of this population is the availability of civil records that would allow reconstructing ascending genealogies and tracing back the mutation to its founder(s). Genealogical information can potentially

increase the power of linkage disequilibrium analysis in fine mapping of the HED locus as demonstrated by Brais *et al.* (in press) in a genetic study of a dominant disease called oculopharyngeal muscular dystrophy in the French Canadian population.

REFERENCES

Ahmad, W., Irvine, A. D., Lam, H., Buckley, C., Bingham, E. A., Panteleyev, A. A., Ahmad, M., McGrath, J. A., and Christiano, A. M. (1998a). A missense mutation in the zinc-finger domain of the human hairless gene underlies congenital atrichia in a family of Irish travellers. *Am. J. Hum. Genet.* **63:** 984-991.

Ahmad, W., ul Haque, M. F., Brancolini, V., Tsou, H. C., ul Haque, S., Lam, H., Aita, V. M.; Owen, Jason; deBlaquiere, M., Frank, J., Cserhalmi-Friedman, P. B., Leask, A., McGrath, J. A., Peacocke, M., Ahmad, M., Ott, J., and Christiano, A. M. (1998b). Alopecia Universalis associated with a mutation in the Human hairless gene. *Science* **279:** 420-724.

Antonarakis, S. E. (1994). Genome linkage scanning: systematic or intelligent? *Nature Genet.* **8:** 211-212.

Brais, B., Bouchard, J. P., Jomphe, M., Desjardins, B., Dubé, M. -P., Gosselin, F., Xie, Y. -G., Tomé, F. M. S., Fardeau, M., Brunet, G., Samson, F., Mathieu, J., Prévost, C., Duranceau, A., Codère, F., Pericak-Vance, M., Bulman, D. E., Tapscott, S., Feasby, T., Munsat, T. L., Pilarski, R., Morgan, K., and Rouleau, G.A.. When genetics and history converge: the fine-mapping and North American introduction and diffusion of the French Canadian oculopharyngeal muscular dystrophy mutation. *Am. J. Hum. Genet.* In press.

Clouston, H. R. (1929) Hereditary ectodermal dystrophy. *Can. Med. Ass. J.* **21:** 18-31.

Collins, F. S. (1995). Positional cloning moves from perditional to traditional. *Nat. Genet.* **9:** 347-50.

Cox, D. R., Burmeister, M., Price, E. R., Kim, S., and Myers, R. M. (1990). Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. *Science* **250:** 245-250.

Craig, J. M., and Bickmore, W. A. (1994). The distribution of CpG islands in mammalian chromosomes. *Nat. Genet.* **7:** 376-82.

Der Kaloustian, V. M., and Kurban, A. K. (1979). In *Genetic diseases of the skin.* Springer-Verlag, pp. 109-114.

Devlin, B., Risch, N., and Roeder, K. (1996). Disequilibrium mapping: composite likelihood for pairwise disequilibrium. *Genomics* **36:** 1-16.

Emonet, N., Michaille, J. J., and Dhouailly, D. (1997). Isolation and characterization of genomic clones of human sequences presumably coding for hair cysteine-rich proteins. *J. Dermatol. Sci.* **14:** 1-11.

Escobar, V., Goldblatt, L. I., Bixler, D., and Weaver, D. (1983). Clouston sydrome: an ultrastructural study. *Clin. Genet.* **24:** 140-146.

Giraud, F., Mattei, J. F., Rolland, M., Ghiglione, C., Pommier de Santi, P., and Sudan, N. (1977). La dysplasie ectodermique de type Clouston. *Arch. Franç. Péd.* **34:** 982-993.

Gold, R. J. M., and Scriver, C. R. (1972). Properties of hair keratin in an autosomal dominant form of ectodermal dysplasia. *Am. J. Hum. Genet.* **24:** 549-561.

Greig, G. M., Warburton, P. E., and Willard, H. F. (1993). Organization and evolution of an alpha satellite DNA subset shared by human chromosome 13 and 21. *J. Mol. Evol.* **37:** 464-475.

Guilford, P., Dode, C., Crozet, F., Blanchard, S., Chaib, H., Levilliers, J., Levi-Acobas, F., Weil, D., Weissenbach, J., Cohen, D., Le Paslier, D., Kaplan, J-C., and Petit, C. (1995). A YAC contig and an EST map in the pericentromeric region of chromosome 13 surrounding the loci for neurosensory nonsydromic deafness (DFNB1 and DFNA3) and limb-girdle muscular dystrophy type 2C (LGMD2C). *Genomics* **29:** 163-169.

Hästbacka, J., de la Chapelle, A., Kaitila, I., Sistonen, P., Weaver, A., and Lander, E. (1992). Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. *Nature Genet.* **2:** 204-211.

Hayflick. S. J., Taylor, T., McKinnon, W., Guttmacher, A. E., Litt, M., and Zonana, J. (1996). Clouston syndrome (hidrotic ectodermal dysplasia) is not linked to keratin gene clusters on chromosomes 12 and 17. *J. Invest. Dermatol.* **107:** 11-14.

Hennies, H. C., Kuster, W., Mischke, D., and Reis, A. (1995). Localization of a locus for the striated form of palmoplantar keratoderma to chromosome 18q near the desmosomal cadherin gene cluster. *Hum. Mol. Genet.* **4:** 1015-1020.

Holbrook, K.A. (1988). Structural abnormalities of the epidermally derived appendages in skin from patients with ectodermal dysplasia: insights into developmental errors. *Birth Defects: Original Article Series.* 24: 15-44.

Holbrook, K. A., Smith, L. T., Kaplan, E. D., Minami, S. A., Hebert, G. P., and Underwood, R. A. (1993). Expression of morphogens during human follicle

development in vivo and a model for studying follicle morphogenesis in vitro. *J. Invest. Dermatol.* **101(1 Suppl):** 39S-49S.

Kaplan, N. L., Hill, W. G., and Weir, B. S. (1995). Likelihood methods for locating disease genes in nonequilibrium populations. *Am. J. Hum. Genet.* **56:** 18-32.

Kelsell, D. P., Dunlop, J., Stevens, H. P., Lench, N. J., Liang, J. N., Parry, G., Mueller, R. F., and Leigh, I. M. (1997). Connexin 26 mutations in hereditary non-syndromic sensorineural deafness. *Nature* **387:** 80-83.

Kere, J., Srivastava, A. K., Montonen, O., Zonana, J., Thomas, N., Ferguson, B., Munoz, F., Morgan, D., Clarke, A., Baybayan, P., Chen, E. Y., Ezer, S., Saariallo-Kere, U., de la Chapelle, A., and Schlessinger, D. (1996). X-linked anhidrotic (hypohidrotic) ectodermal dysplasia is caused by mutation in a novel transmembrane protein. *Nat. Genet.* **13:** 409-416.

Kibar, Z., Dubé, M. –P., Powell, J., McCuaïg, C., Hayflick, S. J., Zonana, J., Hovnanian, A., Radhakrishna, U., Antonarakis, S. E., Benohanian, A., Sheeran, A. D., Stephan, M. L., Gosselin, R., Kelsell, D. P., Christianson, A. L., Fraser, F. C., Der Kaloustian, V. M., and Rouleau, G. A. Clouston hidrotic ectodermal dysplasia (HED): genetic homogeneity, presence of a founder effect in the French Canadian population and fine genetic mapping. Submitted.

Lander, E. S. (1996). The new genomics: global views of biology. *Science* **274:** 536-539.

Lander, E., and Green, P. (1987). Construction of multilocus genetic linkage maps in humans. *Proc. Natl. Acad. Sci. USA* **84:** 2363-2367.

Mackay, D., Ionides, A., Kibar, Z., Rouleau, G., Berry, V., Moore, A., Shiels, A., and Bhattacharya, S. (1999). Connexin 46 mutations in autosomal dominant congenital cataract. *Am. J. Hum. Genet.* **60:** 1474-1478.

Mackinnon, P. J., Powell, B. C., Rogers, G. E., Baker, E. G., Mackinnon, R. N., Hyland, V. J., Callen, D. F., and Sutherland, G. R. (1991) An ultra-high sulfur keratin gene of the human hair cuticle is located at 11q13 and cross-hybridizes with sequences at 11p15. *Mammalian Genome* **1:** 53-56.

Matise, T. C., Perlin, M., and Chakravarti, A. (1994). Automated construction of genetic linkage maps using an expert system (MultiMap): a human genome linkage map. *Nature Genet.* **6:** 384-390.

Mattei, M. G., Penault-Llorca, F., Coulier, F., and Birnbaum, D. (1995). The human FGF9 gene maps to chromosomal region 13q11-q12. *Genomics* **29:** 811-812.

Mignon, C., Fromaget, C., Mattei, M. G., Gros, D., Yamasaki, H., and Mesnil, M. (1996). Assignment of connexin 26 (GJB2) and 46 (GJA3) genes to human chromosome 13q11-->q12 and mouse chromosome 14D1-E1 by in situ hybridization. *Cytogenet. Cell Genet.* **72:** 185-186.

Monreal, A. W., Zonana, J., and Ferguson, B. (1998). Identification of a new splice form of the eDA1 gene permits detection of nearly all X-linked hypohidrotic ectodermal dysplasia mutations. *Am. J. Hum. Genet.* **63:** 380-389.

Moll, I., Heid, H. W., Franke, W. W., and Moll, R. (1988). Patterns of expression of trichocytic and epithelial cytokeratins in mammalian tissues. III. Hair and nail formation during human fetal development. *Differentiation* **39:** 167-184.

Montonen, O., Ezer, S., Saarialho-Kere, U. K., Herva, R., Karjalainen-Lindsberg, M. L., Kaitila, I., Schlessinger, D., Srivastava, A. K., Thesleff, I., and Kere, J. (1998). The gene defective in anhidrotic ectodermal dysplasia is expressed in the developing epithelium, neuroectoderm, thymus, and bone. *J. Histochem. Cytochem.* **46:** 281-289.

Powell, B. C, Nesci, A., and Rogers, G. E. (1991). Regulation of keratin gene expression in hair follicle differentiation. *Ann. N. Y. Acad. Sci.* **642:** 1-20.

Radhakrishma, U., Blouin, J.-L., Mehenni, H., Mehta, T. Y., Sheth, F. J., Sheth, J. J., Solanki, J. V., and Antonarakis S.E. (1997). The gene for autosomal dominant hidrotic ectodermal dysplasia (Clouston syndrome) in a large Indian family maps to the 13q11-q12.1 pericentromeric region. *Am. J. Med. Genet.* **71:** 80-86.

Reiter, A., Sohal, J., Kulkarni, S., Chase, A., Macdonald, D. H., Aguiar, R. C., Goncalves, C., Hernandez, J. M., Jennings, B. A., Goldman, J. M., and Cross, N. C. (1998). Consistent fusion of ZNF198 to the fibroblast growth factor receptor-1 in the t(8;13)(p11;q12) myeloproliferative syndrome. *Blood* **92:** 1735-1742.

Saffer, J. D., Thurston, S. J., Annarella, M. B., and Compton, J. G. (1990). Localization of the gene for the trans-acting transcription factor Sp1 to the distal end of mouse chromosome 15. *Genomics* **8:** 571-574.

Schlessinger, D. (1990). Yeast artificial chromosomes: tools for mapping and analysis of complex genomes. *Trends Genet.* **6:** 248, 255-258.

Simrak, D., Cowley, C. M., Buxton, R. S., and Arnemann, J. (1995). Tandem arrangement of the closely linked desmoglein genes on human chromosome 18. *Genomics* **25:** 591-594.

Stenn, K. S., Combates, N. J., Eilertsen, K. J., Gordon, J. S., Pardinas J. R., Parimoo, S., and Prouty, S.M.(1996). Hair follicle growth controls. *Dermatol Clin.* **14:** 543-558.

Taylor, T.D., Hayflick, S. J., McKinnon, W., Guttmacher, A. E., Hovnanian, A., Litt, M., and Zonana, J. (1998). Confirmation of linkage of Clouston syndrome (hidrotic ectodermal dysplasia) to 13q11-q12.1 with evidence for multiple independent mutations. *J. Invest. Dermatol.* **111:** 83-85.

Tenenhouse, H. S., and Gold, R. J. M. (1976). Loss of a homologous group of proteins in a dominantly inherited ectodermal malformation. *Biochem .J.* **159:** 149-157.

Terwilliger, J. D. (1995). A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am. J. Hum. Genet.* **56:** 777-787.

Van Camp, G., Cruts, M., Backhovens, H., Wehnert, A., and Van Broeckhoven, C. (1992). Unique sequence homology in the pericentromeric regions of the long arms of chromosomes 13 and 21. *Genomics* **12:** 158-160.

Wang, M., Kim, I. G., Steinert, P. M., and McBride, O. W. (1994) Assignment of the human transglutaminase 2 (TGM2) and transglutaminase 3 (TGM3) genes to chromosome 20q11.2. *Genomics* **23:** 721-722.

Williams, M., and Fraser, F. C. (1967). Hydrotic ectodermal dysplasia-Clouston's family revisited. *Can. Med. Ass. J.* 96: 36-38.

Wilkie, A. O., Morriss-Kay, G. M., Jones, E. Y., and Heath, J. K. (1995). Functions of fibroblast growth factors and their receptors. *Curr. Biol.* **5:** 500-507.

Winter, H., Rogers, M. A., Langbein, L., Stevens, H. P., Leigh, I. M., Labrèze, C., Roul, S., Taieb, A., Krieg, T., and Schweizer, J. (1997b). Mutations in the hair cortex keratin hHb6 cause the inherited hair disease monilethrix. *Nature Genet.* 16: 372-374.

APPENDIX