

Molecular genetics of the Finnish disease heritage

Leena Peltonen^{1,2,+}, Anu Jalanko¹ and Teppo Varilo¹

¹Department of Human Molecular Genetics, National Public Health Institute, Helsinki, Finland and ²Department of Human Genetics, UCLA School of Medicine, 695 Charles E. Young Drive South, Los Angeles, CA 90095-7088, USA

Received June 7, 1999; Accepted June 16, 1999

Finland, located at the edge of the inhabitable world, is one of the best-studied genetic isolates. The characteristic features of population isolates—founder effect, genetic drift and isolation—have, over the centuries, shaped the gene pool of the Finns. Finnish diseases have been a target of extensive genetic research and the majority of some 35 disease genes enriched in this population have been identified; the molecular and cellular consequences of disease mutations are currently being characterized. Special strategies taking advantage of linkage disequilibrium have been efficiently used in the initial mapping and restriction of Finnish disease loci and this has stimulated development of novel statistical approaches in the disease gene hunt. Identification of mutated genes has provided tools for detailed analyses of molecular pathogenesis in Finnish diseases, many of which reveal a distinct tissue specificity of clinical phenotype. Often these studies have not only clarified the molecular detail of Finnish diseases, but also provided novel information on biological processes and metabolic pathways essential for normal development and function of human cells and tissues.

THE POPULATION OF FINLAND

The dual theory of the inhabitation of Finland, supported by recent analyses of Y chromosome haplotypes, assumes an early migratory wave of eastern Uralic speakers some 4000 years ago with a distinct effect on today's Finnish gene pool. The majority of the genes of today's Finnish population is thought to originate from later small founder populations which in the beginning of the first millennium immigrated from the south over the Gulf of Finland (1–6). This migration of Indo-European speakers some 2000 years ago was also accompanied by language replacement (Fig. 1A; 7).

Since there is no archaeological, historical, linguistic or genetic evidence for any single major migratory movement to Finland, it is actually more likely that Finland has been inhabited without interruption from the last glacial period, constantly receiving small but significant immigrant groups, mainly from the south and east, but also from the west, throughout prehistoric times.

Both paternally inherited Y chromosome haplotypes and maternally inherited mitochondrial sequences show an exceptional decrease in the genetic diversity of the Finns when compared with other European populations (7). The reasons for the relative isolation of the Finnish population throughout the centuries are mostly geographical. The Baltic Sea surrounds the country to the south and west and the Arctic Ocean to the north. The geopolitical position between Sweden and Russia, between two areas of distinct culture, language and religion, also had an impact on the isolation of Finland. For centuries, only a narrow strip of land in the coastal areas of the south and southwest was populated and the number of settlers must have been small (7). In the 12th century the population of Finland

was only ~50 000. The population fluctuated and reached 250 000 in the 16th century, habitation still being heavily concentrated in the coastal areas.

An internal migration movement began in the 16th century from a small southeastern area to the middle, western and finally northern and eastern parts of the country (Fig. 1B). With increasing population there was pressure to cultivate more land. Another main reason for internal immigration was to avoid the increasing taxation of the Swedish Crown, which had developed a good administrative infrastructure, even reaching into the backlands of Finland. Especially during the regime of the Swedish King Gustavus of Vasa in 1523–1560, inhabitation of wilderness was highly favored. The era of his regime also resulted in the creation of church records, a major resource for later genetic studies of the Finns. These records report births and deaths, marriages and moves of families and provide a highly reliable source of genealogical information, especially since >90% of the population still belongs to the Evangelic Lutheran State Church.

Within a century the inhabited land area of Finland doubled. As the settlers continued to the north, the Saami native minority (Lapps) was forced to move towards the Arctic Ocean. By the end of the 17th century almost the entire land area of Finland was sparsely but permanently populated by Finns. The great famine during the years 1696–1698 and the epidemics that followed were a setback, as one-third of the population of 400 000 was lost. Since then, the Finnish population has grown relatively rapidly from 250 000 at the beginning of the 18th century to its present 5 100 000 (Fig. 2). The internal migration in the 16th century resulted in isolated rural populations that remained surprisingly stable until the Second World War and industrialization.

⁺To whom correspondence should be addressed at: Department of Human Genetics, UCLA School of Medicine, 695 Charles E. Young Drive South, Los Angeles, CA 90095-7088, USA. Tel: +1 310 794 5631; Fax: +1 310 794 5446; Email: lpeltonen@mednet.ucla.edu

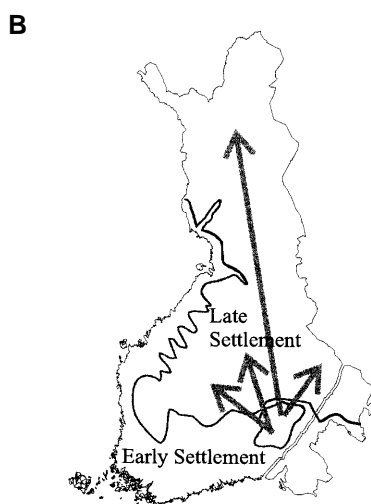
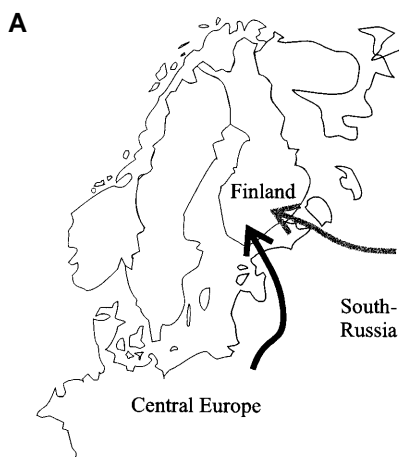


Figure 1. (A) The main migrations of the dual theory of inhabitation of Finland. (B) The internal migration movement of the 1500s resulting in regional subsisolates.

THE CONCEPT OF THE FINNISH DISEASE HERITAGE

Twenty rare inherited disorders more prevalent in Finland than in other populations were initially described in 1973 and this introduced the concept of the Finnish disease heritage (8). Today this group includes >30 typically recessive diseases (5,9) and new diseases with well-defined clinical phenotypes are still being added (10).

The Finnish disease heritage has its origins in the special population history of Finland, briefly described above. The small number of original founders, followed by isolation, rapid expansion and the sampling of small immigrant groups from the main population, allowed the founder effect and genetic drift to mold the gene pool, especially in the regional subsisolates (2,3,8). The frequency of some rare diseases increased, creating the Finnish disease heritage, whereas some disease alleles became almost non-existent (e.g. phenylketonuria, galactosemia, maple syrup disease and cystic fibrosis). In Finnish subsisolates a small number of founders unavoidably

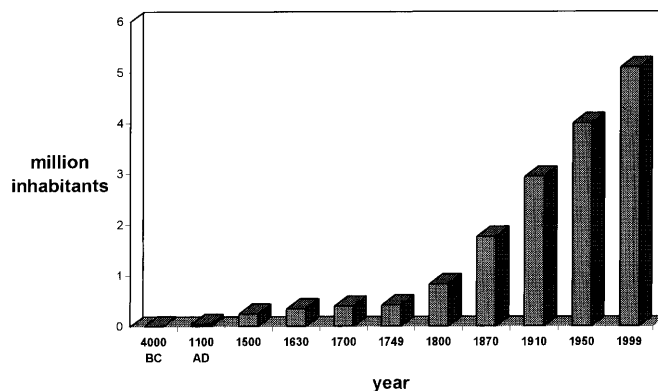


Figure 2. The population growth of Finland.

resulted in consanguineous marriages, although relationships between individuals were typically several generations old and unknown to them. This random inbreeding increased the local incidence of rare recessive disorders and in some Finnish diseases a regional clustering of cases can still be observed, reflecting the regional origin of the founders. Good examples are the cluster of families with progressive epilepsy with mental retardation (EPMR) (11) on the eastern border and the similar cluster of the variant form of late infantile neuronal ceroid lipofuscinosis (vLINCL) cases on the western coastline (Fig. 3; 12). It should be emphasized that in most Finnish diseases solid evidence of the true relationship between affecteds cannot be obtained. The church records reach back to 1634 and the land tax registers to the 1540s and any claim for older relationships can only be based on the origin of founders in nearby or the same communities.

IDENTIFICATION OF FINNISH DISEASE LOCI

The Finnish disease heritage has been the subject of intensive genetic and molecular research over the past 10 years. The loci for 32 of these diseases have been localized and the causative genes for 17 diseases have been isolated, many of them by positional cloning (Table 1) and, according to expectations, all show extreme locus and allelic homogeneity. The molecular data have unequivocally proved the isolation of and significant founder effect in this population. None of these diseases shows locus heterogeneity within Finland, although some, like neonatally lethal Meckel syndrome or late onset frontal lobe dementia with bone cysts (PLO-SL), demonstrate locus heterogeneity outside Finland (40,44). Furthermore, for all the cloned Finnish disease genes, one major mutation is present in $\geq 70\%$ of the disease chromosomes, representing as much as 98% of some disease alleles (Table 1; 15,23,45). Since some of diseases are spread throughout the country, they must be the result of an old mutation event carried to Finland with an early immigrant. One major mutation still identified in all diseases is evidence for effective isolation of the population for the past 2000 years.

As a direct consequence of one major founder mutation, linkage disequilibrium (LD), detected with multiple polymorphic markers flanking the disease locus, is obvious in all alleles of the Finnish disease heritage. The genetic interval showing some degree of LD with individual markers at a particular chromosomal region varies considerably in different disease

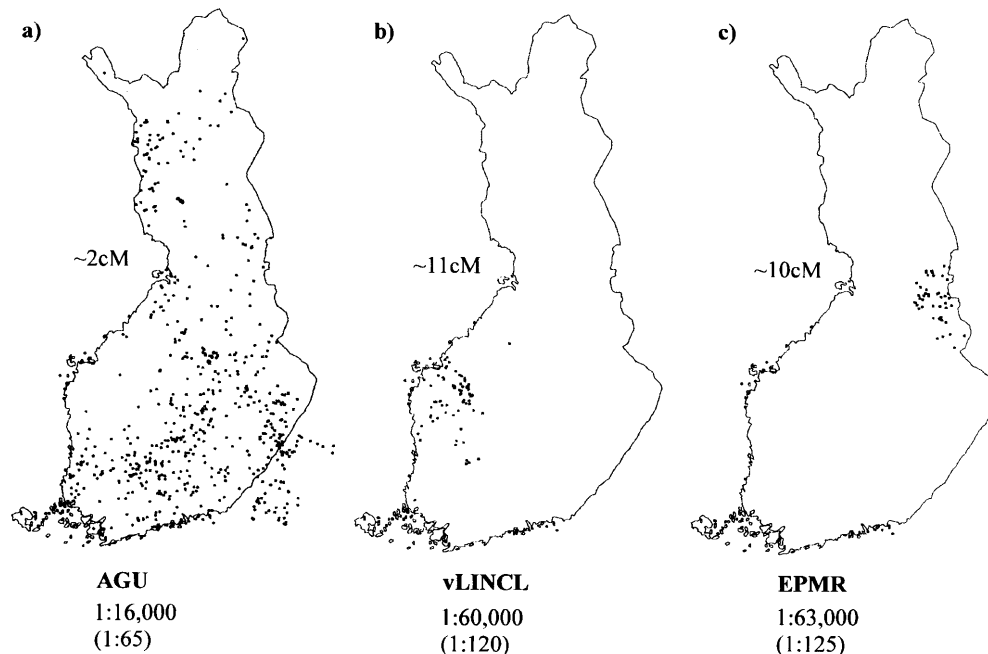


Figure 3. The birthplaces of the grandparents of the patients with AGU, vLINCL and EPMR (R. Norio, 1999, unpublished data). The genetic intervals revealing LD in disease chromosomes reflecting ages of major mutations are given in centiMorgans (cM) (11,12,15). Numbers indicate incidences in Finland. Numbers in parentheses indicate carrier frequencies.

alleles (Table 2). The longer the observed LD interval, the more remarkable is the detectable clustering of grandparents of patients and the more recent the introduction of the mutation into the population (Fig. 3; 3). However, the genetic properties of chromosomal regions and the different demographic histories of Finnish subpopulations also have an impact on the size of the LD interval. The extremes of LD intervals are 13 cM observed in congenital chloride diarrhea and 11 cM in vLINCL chromosomes (12,46).

The first mapping efforts of Finnish diseases were completed using RFLP markers and traditional linkage analyses in a relatively large number of families. This is exemplified by the locus positioning of diastrophic dysplasia (DTD) and infantile neuronal ceroid lipofuscinosis (INCL) mapped with 15 (DTD) and 13 (INCL) families with two or more affected children (47,48). When rarer or neonatally lethal diseases became the target of mapping efforts, special LD-based strategies were applied and genome scans relying on monitoring of genotypes among affecteds or homozygosity of marker genotypes were successfully used. Initial positioning of the locus for the infantile form of spinocerebellar ataxia (IOSCA) was performed using DNA samples of only four patients to search for shared alleles of multiallelic markers (35). Two cases were from the same sibship with the parents being first cousins and two from two sibships with affecteds being separated by seven meioses. A sparse marker set of only 218 markers was used (average marker interval 20 cM) in the initial scan of these four patients and second stage genotyping of all families and segregation-based linkage analyses were only performed for those markers revealing sharing of the alleles. The benefits in the saving of labor and money are obvious; the disease locus was positioned using only a total of 1000 genotypings instead of the >10 000

needed for the traditional, linkage-based strategy in more mixed populations.

Another example of the power of LD is the recent genome scan to identify the locus for a new lethal neonatal metabolic syndrome (10). The initial genome scan was carried out using only four affected sibpairs and searching for markers revealing IBS genotypes in affecteds. Consequent linkage analysis in the small family material consisting of only 20 affected and 16 non-affected chromosomes resulted in only suggestive pairwise lod scores and failed to provide unequivocal evidence for the locus position. However, the linkage disequilibrium ($P < 0.0001$) initially observed with three markers provided proof for the locus assignment. The multipoint linkage disequilibrium analysis (DISMULT) (49) combining information from four markers revealing LD resulted in an impressive lod score of 30.9 and confirmed the locus assignment to 2q33–37 (34).

FINE MAPPING OF FINNISH DISEASE LOCI

The concept of a restricted number of founders and rapid expansion of a population in relative isolation has similarities to the exponential growth of bacterial populations. This motivated Hästbacka *et al.* (47) to adapt bacterial genetics to interpret the LD information in a quantitative fashion using the Luria and Delbrück formula (50) originally designed to estimate mutation rate in rapidly growing bacterial cultures. The formulae allow estimation of the likely mean and standard deviation of the recombination fraction between a disease locus and a marker, the degree of the homogeneity of disease alleles and the mutation rate for genetic markers. The success of this strategy in the evaluation of the distance between a disease locus and a close marker was best demonstrated by the isolation of a diastrophic dysplasia gene precisely at the dis-

Table 1. Identified gene defects and mapped loci for diseases enriched in Finland

| Disease (OMIM no.) | Defective protein | Locus | Major mutation occurrence in Finland (%) | Reference |
|---|---|-------|--|--|
| APECED (240300) | Novel nuclear protein | | 82 | 13,14 |
| Aspartylglucosaminuria (AGU, 208400) | Aspartylglucosaminidase | | 98 | 15 |
| Choroideremia (CHM, 303100) | Rab geranylgeranyl transferase | | n.d. | 16 |
| Congenital chloride diarrhea (CCD, 214700) | Product of the gene down regulated in adenoma | | 100 | 17 |
| Congenital nephrosis (CNF, 256300) | Nephrin | | 78 | 18 |
| Diastrophic dysplasia (DTD, 222600) | Sulfate transporter | | 90 | 19 |
| Familial amyloidosis, Finnish type (FAF, 105120) | Gelsolin | | 100 | 20 |
| Gyrate atrophy of choroid and retina (HOGA, 258870) | Ornithine γ -aminotransferase | | 85 | 21 |
| Hypergonadotrophic ovarian dysgenesis (ODG1, 233300) | Follicle stimulating hormone receptor | | 100 | 22 |
| Infantile neuronal ceroid lipofuscinosis (INCL, 256730) | Palmitoyl protein thioesterase | | 98 | 23 |
| Lysinuric protein intolerance (LPI, 222700) | L-Amino acid transporter | | 100 | 24,25 |
| Non-ketotic hyperglycinemia (NKH, 238300) | Glycine cleavage system; protein P | | 70 | 26 |
| Progressive epilepsy with mental retardation (EPMR, 600143) | Novel transmembrane protein | | n.d. | A.-E. Lehesjoki <i>et al.</i> , unpublished data |
| Progressive myoclonus epilepsy (PME, 254800) | Cystatin B | | 96 | 27 |
| Retinoschisis (RS, 312700) | XLRS1 | | 70 | 28,101 |
| Sialic acid storage disease (SIASD, 268740) | Novel transporter | | 94 | F.Verheijen <i>et al.</i> , unpublished data |
| Finnish variant of late infantile neuronal ceroid lipofuscinosis (vLINCL, 256731) | Novel membrane protein | | 94 | 29 |
| Cartilage-hair hypoplasia (CHH, 250250) | | 9p | | 30 |
| Cohen syndrome (COH1, 216550) | | 8q | | 31 |
| Congenital lactase deficiency (CLD, 223000) | | 2q | | 32 |
| Cornea plana congenita (CNA2, 217300) | | 12q | | 33 |
| Fetal metabolic syndrome with iron accumulation (603358) | | 2q | | 34 |
| Hydroletharus (236680) | | 11q | | 103 |
| Infantile onset spinocerebellar ataxia (IOSCA, 271245) | | 10q | | 35 |
| Lethal congenital contracture-syndrome (LCCS, 253310) | | 9q | | 36 |
| Meckel syndrome (MKS, 249000) | | 17q | | 37 |
| Mulibrey nanism (MUL, 253250) | | 17q | | 38 |
| Muscle-eye-brain disease (MEB, 253280) | | 1q | | 39 |
| Presenile frontal dementia with bone cysts (PLO-SL, 221770) | | 19q | | 40 |
| Selective intestinal malabsorption of vitamin B-12 (MGA1, 261100) | | 10p | | 41 |
| Tibial muscular dystrophy (TMD, 600334) | | 2q | | 42 |
| Usher syndrome, type III (USH, 276902) | | 3q | | 43 |

tance approximated in the original linkage map. The gene was predicted to lie 0.06 cM (or 60 kb) proximal of the best marker and was finally isolated 70 kb proximal of it (19,47). Reverse use of the Luria–Delbrück formula has also been adapted to estimate the age of the mutation. For congenital chloride diarrhea the estimate was 20 generations, in good agreement with the population history of the spread of this mutation in an expanding subpopulation in Eastern Finland starting 400 years ago (17). Monitoring the high LD in disease alleles has targeted the search for disease genes to a highly restricted DNA region in several successful positional cloning projects, including cloning of nephrin, a gene mutated in congenital nephrosis, and AIRE, a novel gene mutated in a multisymptomatic autoimmune disease, APECED (13,18).

Monitoring an ancestral haplotype by fine mapping of chromosome regions showing linkage or LD has also helped in locus restriction in all Finnish disease alleles. A marker set providing 0.1–1 cM resolution has without exception revealed one major shared haplotype among affecteds and greatly facilitated multiple positional cloning efforts of Finnish disease genes. It should be emphasized that, as described earlier, although LD can be observed over wide genetic intervals, identical haplotypes of disease alleles in affecteds from different families are systematically found only across highly restricted DNA regions.

Two examples of powerful restriction of the critical DNA region based on a shared haplotype are progressive myoclonus epilepsy (EPM1; OMIM 254800) and PLO-SL (OMIM

Table 2. The interval at which linkage disequilibrium is reported in some Finnish diseases

| Disorder | Gene location | LD interval (cM) |
|---|---------------|------------------|
| Autoimmune polyendocrinopathy–candidiasis–ectodermal dystrophy (APECED) | 21q22.3 | 3 |
| Cartilage–hair hypoplasia (CHH) | 9p21–p13 | 3 |
| Choroideremia | Xq21 | 9 |
| Congenital chloride diarrhea (CCD) | 7p31 | 13 |
| Congenital nephrosis (CNF) | 19q12–q13.1 | 3 |
| Cornea plana congenita | 12q21 | 2.5 |
| Diastrophic dysplasia (DTD) | 5q31–q34 | 2 |
| Familial amyloidosis, Finnish type (FAF) | 9q33 | 3.5 |
| Infantile neuronal ceroid-lipofuscinosis (INCL) | 1p32 | 2.5 |
| Infantile onset spinocerebellar ataxia (IOSCA) | 10q23.3–q24.1 | 5 |
| Northern epilepsy (Kainuu epilepsy) | 8ptel. | 10 |
| Progressive myoclonus epilepsy (PME) | 21q22 | 5 |
| Retinoschisis | Xp22.2–p22.1 | 10 |
| Salla disease | 6q14–q15 | 10 |
| Usher syndrome, type III (USH3) | 3q21–q25 | 7.5 |
| Variant form of late infantile NCL (vLINCL) | 13q21.1–q32 | 11 |

221770), mentioned earlier. The obligatory recombinations observed in linkage analysis assigned the EPM1 locus to a 7 cM region on chromosome 21q. The haplotype shared among affected alleles in 38 families restricted the critical region to 176 kb (51). Similarly, the PLO-SL locus was, in a genome-wide scan, assigned to the 9 cM region bordered by obligatory recombinations and disease alleles showed LD with markers over a 3 cM interval on chromosome 19q. A shared haplotype was seen only over 150 kb in 14 families (40). These examples demonstrate the power of ancient recombinations in restriction of the critical DNA region to an interval accessible by molecular tools. All reported positional cloning projects on Finnish diseases have efficiently used haplotype-based restriction of the critical chromosomal DNA region, which has provided the basis for successful isolation of the disease gene.

LIGHTNING HITTING THE SAME SPOT TWICE?

One interesting feature in the mapped disease loci or identified causative genes of the Finnish disease heritage is that some loci are, surprisingly, located within the same restricted chromosomal regions. On chromosome 17 two severe diseases affecting multiple tissues (Meckel syndrome and Mulibrey nanism) are positioned within the same 1.2 Mb DNA clone contig (Table 1; 52). The disease-associated alleles carry different haplotypes, although in both diseases one major haplotype is obvious. LD is detected over an interval of 2 cM in Mulibrey nanism and of 1 cM in Meckel alleles, suggesting an earlier introduction of the Meckel mutation into the population. Similarly, on chromosome 19, the nephrin gene, mutated in congenital nephrosis, and the still unknown gene behind PLO-SL are located within the same 150 kb physical contig (Table 1; 53). Again, a totally different haplotype was identified in these two disease alleles and the interval revealing LD indicated different ages for these two mutations. It is intriguing that independent disease mutations enriched in this population have hit so close to each other at historical times separated by tens of generations. Analyses of the full sequence information of these chro-

somal regions could provide some explanation for this interesting feature.

FUNCTIONAL ANALYSES OF FINNISH DISEASE GENES

Thoroughly characterized phenotypes in patients with Finnish diseases provide a profound basis for functional studies of newly identified, often novel disease genes. This again has not only led to an understanding of the disease mechanisms but in multiple cases revealed novel biological information on essential metabolic pathways. Currently, a total of 17 genes of Finnish diseases have been isolated and for most of these some functional data already exist. Some genes have proved to code for essential molecules with highly organ-specific functions. They either show target tissue-specific expression, like nephrin in congenital nephrosis (18,54), chloride transporter in congenital chloride diarrhea (17,100) and FSHR in ovarian dysgenesis (22,56), or encode a component of a metabolic pathway that although present in multiple tissues is crucial for a specific tissue, like DTSD for cartilage in diastrophic dysplasia (19,55) or REP1 in choroideremia. The recently identified gene for an inherited autoimmune disease carries features of a transcription regulatory element (AIRE in APECED) (13,57) but the explanation for multiple tissue symptoms remains to be solved. Ten of the cloned genes in Finnish diseases are involved in neuronal dysfunction. In six of these molecular studies have led to findings of general biological significance and exposed novel details of important metabolic pathways in the CNS. Functional studies of some representative examples are described briefly here.

Aspartylglucosaminuria—exposing new features of lysosomal enzymes

Studies on the molecular mechanisms behind aspartylglucosaminuria (AGU; OMIM 208400) provide a good example how research on a rare disease has resulted in novel informa-

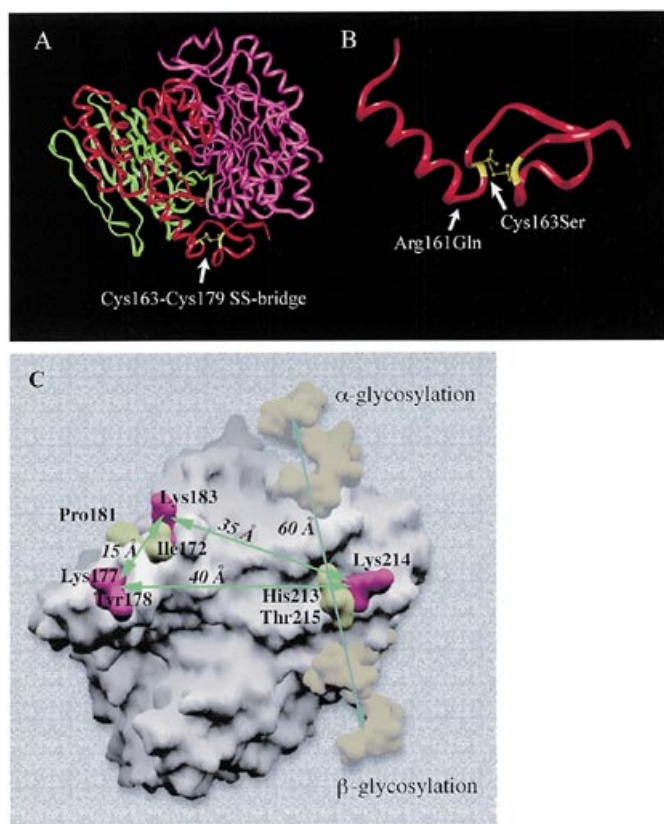


Figure 4. Three-dimensional structure of the AGA enzyme, defective in AGU. (A) The AGU_{Fin} double mutation shown in the three-dimensional structure of the tetrameric ($\alpha\beta\alpha$) AGA molecule. The α -subunit is shown in red and the β -subunit in green, while one $\alpha\beta$ -dimer is purple. (B) Higher magnification of the C-terminus of the α -subunit shows the Cys163Ser substitution resulting from the AGU_{Fin} mutation. The mutation disturbs primary folding of the polypeptide and prevents activation of the AGA molecule, while Arg161Gln represents a neutral polymorphism. (C) The solvent-accessible surface structure of AGA. Three separate regions shown in red on the surface of AGA form the phosphotransferase recognition site. The residues whose substitution reduces phosphorylation of the glycans are shown in purple, while substitution of the light green residues does not affect phosphorylation necessary for lysosomal targeting. The glycans to be phosphorylated are separated by 60 Å.

tion as to the general biology of lysosomal enzymes. AGU results from deficient activity of lysosomal aspartylglucosaminidase (AGA) and the disease results in progressive mental retardation. Following identification of the gene the major disease-causing mutation (AGU_{Fin}) was shown to reside in 98% of disease alleles in Finland. Cellular expression studies of the normal and mutated enzyme provided an understanding of the complex intracellular synthesis, assembly and activation of this heterotetrameric enzyme molecule (15,58–60,64). Crystallographic analyses of the three-dimensional structure of the AGA enzyme revealed a novel catalytic mechanism based on the N-terminal nucleophile and indicated that AGA is the first eukaryotic member of a novel, recently identified enzyme family of N-terminal nucleophile hydrolases (61,62). Detailed structural analyses also led to the characterization of a novel three-dimensional determinant essential for phosphotransferase recognition and lysosomal targeting of AGA, potentially providing a general structural motif for the correct mannose 6-phosphorylation signal of all lysosomal enzymes (Fig. 4; 63).

Knock-out mouse models have replicated the tissue pathology of human disease amazingly well (65,66). Aga^{-/-} mice demonstrate not only the characteristic lysosomal storage vacuoles in CNS neurons but MRI also revealed signs of brain atrophy similar to that seen in older AGU patients. By immunohistochemistry and MRI a subtle delay in myelination was also observed and, similar to the slow clinical course observed in human patients, Aga^{-/-} mice have behavioral symptoms that emerge at an older age (67). The AGU mouse model has further created possibilities to analyze the relationship between myelination and the AGA enzyme and laid the basis for the first therapeutic interventions in AGU (68–70).

Neuronal ceroid lipofuscinoses—a novel class of lysosomal disorders

Neuronal ceroid lipofuscinoses (NCL diseases) constitute a group of severe neurodegenerative diseases all revealing similar tissue pathology with autofluorescent ceroid lipofuscin accumulation bodies as well as hallmark clinical features, including blindness and early death of neocortical neurons. Three of these, i.e. the most severe infantile form, INCL (OMIM 256730), vLINCL (OMIM 256731) and EPMR (CLN8; OMIM 600143), belong to the group of Finnish diseases. The genes behind vLINCL and EPMR were only recently identified (29; A.-E. Lehesjoki *et al.*, unpublished data) and, consequently, functional data exist so far only for INCL. The causative gene for INCL encodes a palmitoyl protein thioesterase (PPT). This finding revealed for the first time the absolute necessity for correct removal of palmitoyl residues from lipid-modified proteins for post-natal survival of neocortical neurons (23). Somewhat surprisingly, PPT was shown to be a lysosomal enzyme, thus classifying INCL as a lysosomal storage disorder (71,72). Ninety-eight percent of Finnish INCL alleles carry the same Arg22Trp mutation, resulting in almost complete failure of transport of the mutant PPT into lysosomes (71). Analyses of the neuronal trafficking of PPT have shown that PPT is transported to the synaptic region in cortical neurons, implying an important role for PPT in synaptic transmission, most probably via interference with membrane recycling, which is extensive in neurons (O. Heinonen *et al.*, unpublished data). Recent data have revealed that other proteins are also defective in NCL diseases; battenin in CLN3 (OMIM 204200) (73) and pepinase in CLN2 (OMIM 204500) (74) represent lysosomal components. An intriguing question emerges as to whether these proteins or the pathways they are involved in interact, since the defects in all of them result in a surprisingly similar tissue pathology, lethal to neocortical neurons although the genes are expressed in all tissues.

Progressive myoclonus epilepsy (Unverricht–Lundborg disease), EPM1—linking minisatellite expansion to apoptosis

Progressive myoclonus epilepsy (PME, Unverricht–Lundborg disease; OMIM 254800) is characterized by tonic-clonic seizures, myoclonic seizures and progressive neurological dysfunction, including dementia and ataxia. Mutations in cystatin B (CSTB) are responsible for EPM1 disease (27), but the disease mechanism is so far unknown. In the majority of Finnish EPM1 patients (96% of disease alleles) the underlying muta-

tion is an unstable minisatellite expansion in the promoter region of CSTB (75,76). In humans EPM1 shows a phenotypic triad consisting of myoclonic seizures, progressive neurological decline and occasional tonic-clonic seizures. Mice lacking CSTB develop myoclonic seizures and ataxia, similar to the symptoms seen in the human disease. The tissue pathology includes loss of cerebral granule cells, which display condensed nuclei, fragmented DNA and other cellular changes characteristic of apoptosis (77).

The defective protein, cystatin B, belongs to a large class of proteins that inhibit cysteine proteases and *in vitro* experiments have shown that cystatin B is a reversible inhibitor of cathepsins B, H, L and S (78,79). Although the cathepsins are localized to lysosomes, cystatin B resides in the cytosol where it probably has a role in the regulation of proteolysis. This suggests that the neuronal pathology of EPM1 results from excessive proteolysis, most harmful to normal neuronal metabolism. The most likely mechanism for neuronal degradation is that the cathepsins, inhibited by cystatin B, directly activate the caspases, thus leading to initiation of apoptosis, and that symptoms of EPM1 are initiated by neuronal apoptosis.

Choroideremia—tissue specificity in the prenylation defect

Choroideremia (OMIM 303100) represents an example of X-linked Finnish diseases characterized by progressive dystrophy of the choroid, retinal pigment epithelium and retina and resulting in blindness. The gene responsible for this disorder was identified by positional cloning in 1990 (80) but only later was the gene product identified as REP-1 (Rab-escort protein) (81,82). Lymphoblasts of choroideremia patients are deficient in Rab geranyl-geranyltransferase, which can be reversed by the addition of REP-1 (83). Rab geranyl-geranyltransferase acts on Rab proteins, which regulate vesicle transport through both the exocytic and endocytic pathways by controlling the assembly of protein complexes involved in vesicle targeting and fusion. The trafficking function of Rab proteins depends upon their binding to cell membranes and this is in turn dependent on lipid modifications. Consequently, modification of Rab proteins by geranyl-geranyltransferase is essential for their function. Rab geranyl-geranyltransferase has a very low affinity for the Rab protein and Rab proteins are recognized only when present in a stable complex with REP (84).

Again, the specificity of the disease in degeneration of the retina and choroid is difficult to account for, since Rab proteins are present in all cells and geranylgeranylation is absolutely required for Rab function. One explanation would be that loss of REP-1 results in defective prenylation of one Rab protein, Rab27 which is expressed in the choroid and retina. As a consequence of the loss of Rab27 function, the choroid epithelial layer in the eye undergoes degeneration leading to loss of vision (83,84).

Familial amyloidosis of the Finnish type—abnormal proteolysis triggering amyloid formation

Familial amyloidosis of the Finnish type (FAF; OMIM 105120), one of two dominant diseases enriched in Finland (Table 1), is a polyneuropathy with lattice dystrophy of the cornea and cranial and peripheral neuropathy (85). The amyloid fibrils in FAF consist of specific peptides proteolytically cleaved from gelsolin and somewhat exceptionally this disease

gene was identified based on biochemical characterization of the tissue amyloid (86). All Finnish FAF patients have a single point mutation in the gelsolin gene changing Asp187 to Asn (20,90). This same mutation is, surprisingly, also observed in Japanese FAF patients on a totally different haplotype, suggesting a mutation hotspot in this particular gene region. World-wide the only other FAF mutation, also hitting the same nucleotide, has been identified in Danish and Czech (Asp187Tyr) (87) patients. This homogeneity of FAF mutations is somewhat surprising for a dominant trait. Gelsolin is an actin-modulating protein which exists in intracellular and secreted forms, both encoded by a single gene on chromosome 9 (88). Gelsolin carrying the Finnish mutation is abnormally proteolytically processed, resulting in secretion of an N-terminally truncated amyloid precursor fragment and this triggers further proteolytic cleavage resulting in amyloid peptides in tissues (89). The capacity for this pathological cleavage by a so far unknown proteolytic activity seems to vary between different cell and tissue types, being most prominent in cells of neuronal origin. This partially explains the tissue specificity of the symptoms and the accumulation of amyloid in cornea and around peripheral nerves (91).

Congenital nephrotic syndrome of the Finnish type—a novel component of the glomerular filter exposed

Congenital nephrotic syndrome of the Finnish type (CNF; OMIM 256300) is characterized by proteinuria, which begins *in utero* (92,93). The NPHS1 gene mutated in CNF encodes a novel protein called nephrin with highly specific expression in the kidney glomerulus. Nephrin is a transmembrane protein of the immunoglobulin superfamily which is expressed in the slit diaphragm of the kidney glomerulus (18), demonstrating an essential role of nephrin in the normal glomerular filtration process (18,54).

Two major mutations were detected in the Finnish population, named Fin_{major} and Fin_{minor}. Sixty-five percent of the patients are homozygous for the Fin_{major} mutation, 8% homozygous for Fin_{minor} and 16% of patients compound heterozygous (18). Based on both this relatively 'low predominance' of the major Finnish mutation and the distribution of the ancestors of the major mutation over the geographical area of Finland, CNF must be one of the oldest representatives of the Finnish disease heritage.

Diastrophic dysplasia and congenital chloride diarrhea—defects in sulfate transporters

Diastrophic dysplasia (DTD; OMIM 222600) is a recessively inherited osteochondrodysplasia caused by mutations in the diastrophic dysplasia sulfate transporter (DTDST) gene on chromosome 5q31 (19). Ninety percent of Finnish DTD alleles carry a splice site mutation resulting in severely reduced mRNA levels (DTDST_{Fin}). A somewhat peculiar finding was that two additional mutations were identified on the same ancestral haplotype found in 95% of Finnish DTD alleles (94). One explanation for this peculiar feature would be that all these DTD mutations arose in the same haplotype, common in the population some 100 generations ago. This haplotype became admixed with another (larger) wave of immigrants not having this haplotype, which became rare in the following generations. Alternatively, this particular haplotype could be more

prone to mutations, although we are not aware of such a phenomenon in any other disease allele.

The DTDST gene encodes a Na⁺-independent sulfate transporter expressed in all tissues. Recent *in vitro* data demonstrate that Na⁺-dependent sulfate transport, an event more crucial for chondrocytes than any other cell type, is dominantly dependent on the DTDST system. Further, undersulfation of proteoglycans impairs the growth response of the cells to fibroblast growth factor and this also contributes to the hallmark feature of diastrophic dysplasia, severely disturbed endochondral bone formation (55).

Clinical presentation of congenital chloride diarrhea (CLD; OMIM 214700) is a lifetime, potentially fatal diarrhea with a high chloride content. The disease results from mutations in a chromosome 7 gene encoding a transmembrane protein also belonging to the sulfate transporter family. All Finnish patients have a deletion of codon 317 resulting in an in-frame deletion of Val. *In vitro* expression in *Xenopus* oocytes demonstrated that both chloride and sulfate ions are transported by this transmembrane protein and this activity is totally lost in the Val-deleted mutant (100). *In situ* data show preferential expression of this gene in intestinal surface epithelium, explaining the involvement of this protein in intestinal symptoms of patients (17).

Ovarian dysgenesis—new clues to infertility

Ovarian dysgenesis resulting in infertility (OMIM 233300) is caused by mutations in the follicle stimulating hormone receptor (FSH-R) (22). The causative gene was identified by the positional candidate gene strategy in Finnish families very soon after the first description of the clinical phenotype (95). FSH-R is a G-protein-coupled transmembrane receptor containing a large extracellular hormone-binding domain. FSH-R has a critical role in reproduction through the control of gonadal development and gamete production (96). In females FSH-R is expressed in the granulosa cells of the ovaries and it controls follicular growth and ovarian steroidogenesis (97). An inactivating mutation in FSH-R (Ala189Val) found in all Finnish disease alleles produces high gonadotropin levels and streaky gonads associated with primary amenorrhea in females. Males with the same mutation display various degrees of spermatogenic failure or absolute infertility (22,98).

FSH-R-deficient mice have been produced and female mice display a severe phenotype with thin uteri and small ovaries and are sterile because of a block in folliculogenesis (56). Mutant males display small testes, partial spermatogenic failure and reduced fertility. These mice will be of major importance in reproduction research since they provide an excellent model in which to analyse the molecular basis of infertility.

Autoimmune polyendocrinopathy—candidiasis—ectodermal dystrophy (APECED)—novel insights into autoimmunity

APECED is perhaps the Finnish disease of widest general interest since APECED patients have multiple autoimmune symptoms appearing throughout their lifespan and this recessive monogenic disease is expected to provide a short-cut to understanding the molecular events involved in autoimmunity. Hallmark symptoms of APECED are multiple autoimmune endocrinopathies, chronic mucocutaneous candidiasis and ectodermal dystrophies. The novel gene behind APECED,

named AIRE, was recently identified and again one major mutation in Finnish APECED patients could be identified in 85% of Finnish disease alleles (13,14). APECED_{Fin major} is a nonsense mutation resulting in a truncated polypeptide chain which is not correctly targeted intracellularly. In two other populations showing exceptionally high incidences of APECED, Iranian Jews and Sardinians, one major mutation is also found in 100 and 92% of disease alleles, respectively.

Several structural features of the APECED protein, including a nuclear targeting signal, high proline content, PHD-type zinc finger domains and a newly described DNA-binding domain called SAND (99), imply nuclear targeting and DNA binding of this protein. Also, *in vitro* expression studies have demonstrated a nuclear localization and a capacity for transcriptional activation of the AIRE protein (57,102; P. Björns *et al.*, unpublished data). This peculiar protein is found in mature thymocytes and peripheral lymphocytes, suggesting that the biological function could be involved in tolerance development in cellular immunity.

COMMON LESSONS FROM RARE DISEASES

Genetic and molecular studies of Finnish diseases have provided several lessons, important for disease locus mapping and identification generally. Special strategies and LD-based shortcuts in both the initial positioning and the fine mapping of disease loci have proved their power in Finnish diseases and often provided experimental evidence for theoretical assumptions. At the molecular level the identification of novel molecules (like nephrin or APECED protein) has provided new tools with which to tackle unknown details of tissue dysfunction or degradation. Finally, the proteins defective in Finnish diseases have often provided novel insights into the significance of a metabolic pathway or suggested interactions between molecules which were not previously known to interact. This is perhaps best exemplified by the molecular dissection of NCL disorders, which have been transformed from a spectrum of slightly varying clinical phenotypes into well-defined molecular entities resulting from dysfunction of lysosomal molecules.

Genetic and molecular studies of Finnish diseases have without doubt brought into the limelight the definitive advantages of genetic isolates as a resource for human genetic research. If wisely exploited using ethically sound research policies and wisely chosen strategies based on sufficient information of the genealogical history, various population isolates will also be of the greatest importance for genetic studies of polygenic and complex diseases. The example of Finland shows how successful research of genetic diseases has been based on well-recorded population histories, the efforts of skillful clinicians and high quality health care. These advantages have produced reliable diagnoses and excellent population and health care registers, but even more importantly a high level of basic trust by the population of genetic research and consequent high participation rates in genetic studies.

REFERENCES

1. Nevanlinna, H.H. (1972) The Finnish population structure. A genetic and genealogical study. *Hereditas*, **71**, 195–236.
2. de la Chapelle, A. (1993) Disease gene mapping in isolated human populations: the example of Finland. *J. Med. Genet.*, **30**, 857–865.

3. Peltonen, L., Pekkarinen, P. and Aaltonen, J. (1995) Messages from an isolate: lessons from the Finnish gene pool. *J. Biol. Chem.*, **376**, 679–704.
4. Lahermo, P., Sajantila, A., Sistonen, P., Lukka, M., Aula, P., Peltonen, L. and Savontaus, M.-L. (1996) The genetic relationship between the Finns and the Finnish Saami (Lapps): analysis of nuclear DNA and mitochondrial mtDNA. *Am. J. Hum. Genet.*, **58**, 1309–1322.
5. de la Chapelle, A. and Wright, F.A. (1998) Linkage disequilibrium mapping in isolated populations: the example of Finland revisited. *Proc. Natl Acad. Sci. USA*, **95**, 12416–12423.
6. Kittles, R.A., Perola, M., Peltonen, L., Bergen, A.W., Aragon, R.A., Virkkunen, M., Linnoila, M., Goldman, D. and Long, J.C. (1998) Dual origins of Finns revealed by Y chromosome haplotype variation. *Am. J. Hum. Genet.*, **62**, 1171–1179.
7. Sajantila, A., Abdel-Halim, S., Savolainen, P., Bauer, K., Gierig, C. and Pääbo, S. (1996) Paternal and maternal DNA linkages reveal a bottleneck in the founding of the Finnish population. *Proc. Natl Acad. Sci. USA*, **93**, 12035–12039.
8. Norio, R., Nevanlinna, H.R. and Perheentupa, J. (1973) Hereditary diseases in Finland: rare flora in rare soul. *Ann. Clin. Res.*, **5**, 109–141.
9. Peltonen, L. (1997) Molecular background of the Finnish disease heritage. *Ann. Med.*, **29**, 553–556.
10. Fellman, V., Rapola, J., Pihko, H., Varilo, T. and Raivio, K.O. (1998) Iron-overload disease in infants involving fetal growth retardation, lactic acidosis, liver haemosiderosis, and aminoaciduria. *Lancet*, **351**, 490–493.
11. Tahvanainen, E., Ranta, S., Hirvasniemi, A., Karila, E., Leisti, J., Sistonen, P., Weissenbach, J., Lehesjoki, A.E. and de la Chapelle, A. (1994) The gene for a recessively inherited human childhood progressive epilepsy with mental retardation maps to the distal short arm of chromosome 8. *Proc. Natl Acad. Sci. USA*, **91**, 7267–7270.
12. Varilo, T., Savukoski, M., Norio, R., Santavuori, P., Peltonen, L. and Järvelä, I. (1996) The age of human mutation: genealogical and linkage disequilibrium analysis of the CLN5 mutation in the Finnish population. *Am. J. Hum. Genet.*, **58**, 506–512.
13. The Finnish German APECED Consortium (1997) An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. *Nature Genet.*, **17**, 399–403.
14. Naganime, K., Peterson, P., Scott, H.S., Kudoh, J., Minoshima, S., Heino, M., Krohn, K.J., Lalot, M.D., Mullis, P.E., Antonarakis, S.E., Kawasaki, K., Asakawa, S., Ito, F. and Shimizu, N. (1997) Positional cloning of the APECED gene. *Nature Genet.*, **17**, 393–398.
15. Ikonen, E., Baumann, M., Grön, K., Syvänen, A.-C., Enomaa, N., Halila, R., Aula, P. and Peltonen, L. (1991) Aspartylglucosaminuria: cDNA encoding human aspartylglucosaminidase and the missense mutation causing the disease. *EMBO J.*, **10**, 51–58.
16. Sankila, E.M., Tolvanen, R., van den Hurk, J.A., Cremers, F.P. and de la Chapelle, A. (1992) Aberrant splicing of the CHM gene is a significant cause of choroideremia. *Nature Genet.*, **1**, 109–113.
17. Höglund, P., Haila, S., Socha, J., Tomaszewski, L., Saarialho-Kere, U., Karjalainen-Lindsberg, M.-L., Airola, K., Holmberg, C., de la Chapelle, A. and Kere, J. (1996) Mutations of the Down-regulated in adenoma (DRA) gene cause congenital chloride diarrhoea. *Nature Genet.*, **14**, 316–319.
18. Kestilä, M., Lenkkeri, U., Männikkö, M., Lamerdin, J., McCready, P., Putaala, H., Ruotsalainen, V., Morita, T., Nissinen, M., Herva, R., Kashtan, C.E., Peltonen, L., Holmberg, C., Olsen, A. and Tryggvason, K. (1998) Positionally cloned gene for a novel glomerular protein—nephrin—is mutated in congenital nephrotic syndrome. *Mol. Cell*, **1**, 575–582.
19. Hästbacka, J., de la Chapelle, A., Mahtani, M.M., Clines, G., Reeve-Daly, M.P., Daly, M., Hamilton, B.A., Kusumi, K., Trivedi, B., Weaver, A. et al. (1994) The diastrophic dysplasia gene encodes a novel sulfate transporter: positional cloning by fine-structure linkage disequilibrium mapping. *Cell*, **78**, 1073–1087.
20. Levy, E., Haltia, M., Fernandez-Madrid, I., Koivunen, O., Ghiso, J., Prelli, F. and Frangione, B. (1990) Mutation in gelsolin gene in Finnish hereditary amyloidosis. *J. Exp. Med.*, **172**, 1865–1867.
21. Mitchell, G., Looney, J., Brody, L., Steel, G., Suchanek, M., Engelhardt, J., Willard, H. and Valle, D. (1988) Human ornithine- δ -aminotransferase: cDNA cloning and analysis of structural gene. *J. Biol. Chem.*, **263**, 14288–14295.
22. Aittomäki, K., Lucena, J.L.D., Pakarinen, P., Sistonen, P., Tapaninen, J., Gromoll, J., Kaskikari, R., Sankila, E.-M., Lehtväslaiho, H., Engel, A.R., Nieschlag, E., Huhtaniemi, I. and de la Chapelle, A. (1995) Mutation in the follicle-stimulating hormone receptor gene causes hereditary hypergonadotropic ovarian failure. *Cell*, **82**, 959–968.
23. Vesa, J., Hellsten, E., Verkruyse, L.A., Camp, L.A., Rapola, J., Santavuori, P., Hofmann, S.L. and Peltonen, L. (1995) Mutations in the palmitoyl protein thioesterase gene causing infantile neuronal ceroid lipofuscinosis. *Nature*, **376**, 584–587.
24. Borsani, G., Bassi, M.T., Sperandio, M.P., De Grandi, A., Buoninconti, A., Riboni, M., Manzoni, M., Incerti, B., Pepe, A., Andria, G., Ballabio, A. and Sebastio, G. (1999) SLC7A7, encoding a putative permease-related protein, is mutated in patients with lysinuric protein intolerance. *Nature Genet.*, **21**, 297–301.
25. Torrents, D., Mykkänen, J., Pineda, M., Feliubadaló, L., Estévez, R., de Cid, R., Sanjurjo, P., Zorzano, A., Nunes, V., Huoponen, K., Reinikainen, A., Simell, O., Savontaus, M.-L., Aula, P. and Palácín, M. (1999) Identification of SLC7A7, encoding y+LAT-1, as the lysinuric protein intolerance gene. *Nature Genet.*, **21**, 293–296.
26. Kure, S., Takayanagi, M., Narisawa, K., Tada, K. and Leisti, J. (1992) Identification of a common mutation in Finnish patients with nonketotic hyperglycinemia. *J. Clin. Invest.*, **90**, 160–164.
27. Pennacchio, L.A., Lehesjoki, A.-E., Stone, N.E., Willour, V.L., Virtaneva, K., Miao, J., D'Amato, E., Ramirez, L., Faham, M., Koskineniemi, M., Warrington, J.A., Norio, R., de la Chapelle, A., Cox, D.R. and Myers, R.M. (1996) Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1). *Science*, **271**, 1731–1734.
28. Sauer, C.G., Gehrig, A., Warneke-Wittstock, R., Marquardt, A., Ewing, C.C., Gibson, A., Lorenz, B., Jurklics, B. and Weber, B.H.F. (1997) Positional cloning of the gene associated with X-linked juvenile retinoschisis. *Nature Genet.*, **17**, 164–170.
29. Savukoski, M., Klockars, T., Holmberg, V., Santavuori, P., Lander, E.S. and Peltonen, L. (1998) CLN5, a novel gene encoding a putative transmembrane protein mutated in the Finnish variant late infantile neuronal ceroid lipofuscinosis (vLINCL). *Nature Genet.*, **19**, 286–288.
30. Sulisalo, T., Sistonen, P., Hästbacka, J., Wadelius, C., Mäkitie, O., de la Chapelle, A. and Kaitila, I. (1993) Cartilage-hair hypoplasia gene assigned to chromosome 9 by linkage analysis. *Nature Genet.*, **3**, 338–341.
31. Tahvanainen, E., Norio, R., Karila, E., Ranta, S., Weissenbach, J., Sistonen, P. and de la Chapelle, A. (1994) Cohen syndrome gene assigned to the long arm of chromosome 8 by linkage analysis. *Nature Genet.*, **7**, 201–204.
32. Järvelä, I., Enattah, N.S., Kokkonen, J., Varilo, T., Savilahti, E. and Peltonen, L. (1998) Assignment of the locus for congenital lactase deficiency to 2q21, in the vicinity of but separate from the lactase-phlorizin hydrolase gene. *Am. J. Hum. Genet.*, **63**, 1078–1085.
33. Tahvanainen, E., Forsius, H., Karila, E., Ranta, S., Eerola, M., Weissenbach, J., Sistonen, P. and de la Chapelle, A. (1995) Cornea plana congenita gene assigned to the long arm of chromosome 12 by linkage analysis. *Genomics*, **26**, 290–293.
34. Visapää, I., Fellman, V., Varilo, T., Palotie, A., Raivio, K. and Peltonen, L. (1998) Assignment of the locus for a new lethal neonatal metabolic syndrome to 2q33–37. *Am. J. Hum. Genet.*, **63**, 1396–1403.
35. Nikali, K., Suomalainen, A., Terwilliger, J., Koskinen, T., Weissenbach, J. and Peltonen, L. (1995) Random search for shared chromosomal regions: the assignment of a new hereditary ataxia locus. *Am. J. Hum. Genet.*, **56**, 1088–1095.
36. Mäkelä-Bengs, P., Järvinen, N., Vuopala, K., Suomalainen, A., Ignatius, J., Sipilä, M., Herva, R., Palotie, A. and Peltonen, L. (1998) Assignment of the disease locus for the lethal congenital contracture syndrome to a restricted region on chromosome 9q34 by genome scan using five affected individuals. *Am. J. Hum. Genet.*, **63**, 506–516.
37. Paavola, P., Salonen, R., Weissenbach, J. and Peltonen, L. (1995) The locus for Meckel syndrome with multiple congenital anomalies maps to chromosome 17q21–q24. *Nature Genet.*, **11**, 213–215.
38. Avela, K., Lipsanen-Nyman, M., Perheentupa, J., Wallgren-Pettersson, C., Marchand, S., Faure, S., Sistonen, P., de la Chapelle, A. and Lehesjoki, A.-E. (1997) Assignment of the mulibrey nanism gene to 17q by linkage and linkage-disequilibrium analysis. *Am. J. Hum. Genet.*, **60**, 896–902.
39. Corman, B., Avela, K., Pihko, H., Santavuori, P., Talim, B., Topaloglu, H., de la Chapelle, A. and Lehesjoki, A.E. (1999) Assignment of the muscle-eye-brain disease gene to 1p32–p34 by linkage analysis and homozygosity mapping. *Am. J. Hum. Genet.*, **64**, 126–135.
40. Pekkarinen, P., Hovatta, I., Hakola, P. et al. (1998) Assignment of the locus for PLO-SL, a frontal-lobe dementia with bone cysts, to 19q13. *Am. J. Hum. Genet.*, **62**, 362–373.

41. Aminoff, M., Tahvanainen, E., Gräsbeck, R., Weissenbach, J., Broch, H. and de la Chapelle, A. (1995) Selective intestinal malabsorption of vitamin B12 displays recessive Mendelian inheritance: assignment of a locus chromosome 10 by linkage. *Am. J. Hum. Genet.*, **57**, 824–831.
42. Haravuori, H., Mäkelä-Bengs, P., Udd, B., Partanen, J., Pulkkinen, L., Somer, H. and Peltonen, L. (1998) Assignment of the tibial muscular dystrophy locus to chromosome 2q31. *Am. J. Hum. Genet.*, **62**, 620–626.
43. Sankila, E.M., Pakarinen, L., Kääriäinen, H., Aittomäki, K., Karjalainen, S., Sistonen, P. and de la Chapelle, A. (1995) Assignment of an Usher syndrome type III (USH3) gene to chromosome 3q. *Hum. Mol. Genet.*, **4**, 93–98.
44. Roume, J., Genin, E., Cormier-Daire, V., Ma, H.W., Mehaye, B., Attie, T., Razavi-Encha, F., Fallet-Bianco, C., Buenerd, A., Clerget-Darpoux, F., Munnich, A. and Le Merrer, M. (1998) A gene for Meckel syndrome maps to chromosome 11q13. *Am. J. Hum. Genet.*, **63**, 1095–1101.
45. Syvänen, A.-C., Ikonen, E., Manninen, T., Bengström, M., Söderlund, H., Aula, P. and Peltonen, L. (1992) Convenient and quantitative determination of the frequency of a mutant allele using solid-phase minisequencing: application to aspartylglucosaminuria in Finland. *Genomics*, **12**, 590–595.
46. Höglund, P., Sistonen, P., Norio, R., Holmberg, C., Dimberg, A., Gustavson, K.H., de la Chapelle, A. and Kere, J. (1995) Fine mapping of the congenital chloride diarrhea gene by linkage disequilibrium. *Am. J. Hum. Genet.*, **57**, 95–102.
47. 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.
48. Järvelä, I., Schleutker, L., Haataja, L., Santavuori, P., Puhakka, L., Manninen, T., Palotie, A., Sandkuijl, L.A., Renlund, M., White, R., Aula, P. and Peltonen, L. (1991) Infantile form of neuronal ceroid lipofuscinosis (CNL1) maps to the short arm of chromosome 1. *Genomics*, **9**, 170–173.
49. 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.
50. Luria, S.E. and Delbrück, M. (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, **28**, 491–511.
51. Lehesjoki, A.-E., Koskiniemi, M., Norio, R., Tirrito, S., Sistonen, P., Lander, E. and de la Chapelle, A. (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.
52. Paavola, P., Avela, K., Horelli-Kuitunen, N., Bärlund, M., Kallioniemi, A., Idänheimo, N., de la Chapelle, A., Palotie, A., Lehesjoki, A.-E. and Peltonen, L. (1999) High-resolution physical and genetic mapping of the critical region for Meckel syndrome and Mulibrey nanism on chromosome 17q22–q23. *Genome Res.*, **9**, 267–276.
53. Pekkarinen, P., Kestilä, M., Paloneva, J., Terwilliger, J., Varilo, T., Järvi, O., Hakola, P. and Peltonen, L. (1998) Fine-scale mapping of a novel dementia gene, PLO-SL, by linkage disequilibrium. *Genomics*, **54**, 307–315.
54. Ruotsalainen, V., Ljungberg, P., Wartiovaara, J., Lenkkeri, U., Kestilä, M., Jalanko, H., Holmberg, C. and Tryggvason, K. (1999) Nephin is located at the slit diaphragm area of glomerular podocytes. *Proc. Natl Acad. Sci. USA*, **96**, 7962–7967.
55. Satoh, H., Susaki, M., Shukunami, C., Iyama, K.-I., Negoro, T. and Hiraki, Y. (1998) Functional analysis of diastrophic dysplasia sulfate transporter; its involvement in growth regulation of chondrocytes mediated by sulfated proteoglycans. *J. Biol. Chem.*, **273**, 12307–12315.
56. Dierich, A., Sairam, M.R., Monaco, L., Fimia, G.M., Gansmuller, A., LeMeur, M. and Sassone-Corsi, P. (1998) Impairing follicle-stimulating hormone (FSH) signaling *in vivo*: targeted disruption of the FSH receptor leads to aberrant gametogenesis and hormonal imbalance. *Proc. Natl Acad. Sci. USA*, **95**, 13612–13617.
57. Björnses, P., Peltto-Huikko, M., Kaukonen, J., Aaltonen, J., Peltonen, L. and Ulmanen, I. (1999) Localization of the APECED protein in distinct nuclear structures. *Hum. Mol. Genet.*, **8**, 259–266.
58. Ikonen, E., Julkunen, I., Tollersrud, O.-K., Kalkkinen, N. and Peltonen, L. (1993) Lysosomal aspartylglucosaminidase is processed to the active subunit complex in the endoplasmic reticulum. *EMBO J.*, **12**, 295–302.
59. Riikonen, A., Tikkanen, R., Jalanko, A. and Peltonen, L. (1995) Immediate interaction between the nascent subunits and two conserved amino acids Trp34 and Thr206 needed for the catalytic activity of aspartylglucosaminidase. *J. Biol. Chem.*, **270**, 4903–4907.
60. Riikonen, A., Rouvinen, J., Tikkanen, R., Julkunen, I., Peltonen, L. and Jalanko, A. (1996) Primary folding of aspartylglucosaminidase: significance of disulfide bridges and evidence of early multimerization. *J. Biol. Chem.*, **271**, 21340–21344.
61. Oinonen, C., Tikkanen, R., Rouvinen, J. and Peltonen, L. (1995) Three-dimensional structure of human lysosomal aspartylglucosaminidase. *Nature Struct. Biol.*, **2**, 1102–1107.
62. Tikkanen, R., Riikonen, A., Oinonen, C., Rouvinen, J. and Peltonen, L. (1996) Functional analyses of active site residues of human lysosomal aspartylglucosaminidase: implications for catalytic mechanism and auto-catalytic activation. *EMBO J.*, **15**, 2954–2960.
63. Tikkanen, R., Peltola, M., Oinonen, C., Rouvinen, J. and Peltonen, L. (1997) Several co-operational binding sites mediate the interaction of a lysosomal enzyme with the phosphotransferase. *EMBO J.*, **16**, 6684–6693.
64. Saarela, J., Laine, M., Tikkanen, R., Oinonen, C., Jalanko, A., Rouvinen, J. and Peltonen, L. (1998) Activation and oligomerization of aspartylglucosaminidase. *J. Biol. Chem.*, **273**, 25320–25328.
65. Jalanko, A., Tenhunen, K., McKinney, C.E., LaMarca, M.E., Rapola, J., Autti, T., Joensuu, R., Manninen, T., Sipilä, I., Ginns, E. and Peltonen, L. (1998) Mice with an aspartylglucosaminuria mutation similar to humans replicate the pathophysiology in patients. *Hum. Mol. Genet.*, **7**, 265–272.
66. Kaartinen, V., Mononen, I., Voncken, J.-V., Noronkoski, T., Gonzalez-Comez, I., Heisterkamp, N. and Groffen, J. (1996) A mouse model for the human lysosomal disease aspartylglucosaminuria. *Nature Med.*, **2**, 1375–1378.
67. Tenhunen, K., Uusitalo, A., Autti, T., Joensuu, R., Kettunen, M., Kauppinen, R.A., Ikonen, S., LaMarca, M.E., Haltia, M., Ginns, E.I., Jalanko, A. and Peltonen, L. (1998) Monitoring the CNS pathology in aspartylglucosaminuria mice. *J. Neuropathol. Exp. Neurol.*, **57**, 1154–1163.
68. Kytälä, A., Heinonen, O., Peltonen, L. and Jalanko, A. (1998) Expression and endocytosis of lysosomal aspartylglucosaminidase in mouse primary neurons. *J. Neurosci.*, **18**, 7750–7756.
69. Peltola, M., Kytälä, A., Heinonen, O., Rapola, J., Paunio, T., Revah, F., Peltonen, L. and Jalanko, A. (1998) Adenovirus-mediated gene transfer results in decreased lysosomal storage in brain and liver of aspartylglucosaminuria (AGU) mouse. *Gene Ther.*, **5**, 1314–1321.
70. Laine, M., Richter, J., Fahlman, C., Rapola, J., Renlund, M., Peltonen, L., Karlsson, S. and Jalanko, A. (1999) Correction of peripheral lysosomal accumulation in mice with aspartylglucosaminuria by bone marrow transplantation. *Exp. Hematol.*, in press.
71. Hellsten, E., Vesa, J., Olkkonen, V.M., Jalanko, A. and Peltonen, L. (1996) Human palmitoyl protein thioesterase: evidence for lysosomal targeting of the enzyme and disturbed cellular routing in infantile neuronal ceroid lipofuscinosis. *EMBO J.*, **15**, 5240–5245.
72. Verkruyse, L.A. and Hofmann, S.L. (1996) Lysosomal targeting of palmitoyl-protein thioesterase. *J. Biol. Chem.*, **271**, 15831–15836.
73. Järvelä, I., Sainio, M., Rantamäki, T., Olkkonen, V.M., Carpen, O., Peltonen, L. and Jalanko, A. (1998) Biosynthesis and intracellular targeting of the CLN3 protein defective in Batten disease. *Hum. Mol. Genet.*, **7**, 85–90.
74. Sleat, D.E., Donnelly, R.J., Lackland, H., Liu, C.-G., Sohar, I., Pullarkat, R. and Lobel, P. (1997) Association of mutations in a lysosomal protein with classical late-infantile neuronal ceroid lipofuscinosis. *Science*, **277**, 1802–1805.
75. Lalioti, M.D. *et al.* (1997) Identification of mutations in cystatin B, the gene responsible for the Unverricht–Lundborg type of Progressive Myoclonus Epilepsy (EPM1). *Am. J. Hum. Genet.*, **60**, 342–351.
76. Virtaneva, K., D'Amato, E., Miao, J., Koskiniemi, M., Norio, R., Avanzini, G., Franceschetti, S., Michelucci, R., Tassinari, C.A., Omer, S., Pennacchio, L.A., Myers, R.M., Dieguez-Lucena, J.L., Krahe, R., de la Chapelle, A. and Lehesjoki, A.-E. (1997) Unstable minisatellite expansion causing recessively inherited myoclonus epilepsy, EPM1. *Nature Genet.*, **15**, 393–396.
77. Pennacchio, L.A., Bouley, D.M., Higgins, K.M., Scott, M.P., Noebels, J.L. and Myers, R.M. (1998) Progressive ataxia, myoclonic epilepsy and cerebellar apoptosis in cystatin B-deficient mice. *Nature Genet.*, **20**, 251–258.
78. Ritonja, A., Machleidt, W. and Barrett, A.J. (1985) Amino acid sequence of the intracellular cysteine proteinase inhibitor cystatin B from human liver. *Biochem. Biophys. Res. Commun.*, **131**, 1187–1192.
79. Barrett, A.J., Fritz, H., Grubb, A., Isemura, S., Järvinen, M., Katunuma, N., Machleidt, W., Muller-Esterl, W., Sasaki, M. and Turk, V. (1986) Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin. *Biochem. J.*, **236**, 312.

80. Cremers, F.P.M., van de Pol, D.J.R., van Kerkhoff, L.P.M., Wieringa, B. and Ropers, H.-H. (1990) Cloning of a gene that is rearranged in patients with choroideraemia. *Nature*, **347**, 674–677.
81. Seabra, M.C., Brown, M.S., Slaughter, C.A., Südhof, T.C. and Goldstein, J.L. (1992) Purification of component A of Rab geranyl-geranyl-transferase: possible identity with choroideremia gene product. *Cell*, **70**, 1049–1057.
82. Seabra, M.C., Goldstein, J.L., Südhof, T.C. and Brown, M.S. (1992) Rab geranylgeranyl transferase: a multisubunit enzyme that prenylates GTP-binding proteins terminating in Cys-X-Cys or CysCys. *J. Biol. Chem.*, **267**, 14497–14503.
83. Seabra, M.C., Brown, M.S. and Goldstein, J.L. (1993) Retinal degeneration in choroideremia: deficiency of Rab geranylgeranyl transferase. *Science*, **259**, 377–381.
84. Andres, D.A., Seabra, M.C., Brown, M.S., Armstrong, S.A., Smeland, T.E., Cremers, F.P. and Goldstein, J.L. (1993) cDNA cloning of component A of Rab geranylgeranyl transferase and demonstration of its role as a Rab escort protein. *Cell*, **73**, 1091–1099.
85. Meretoja, J. (1969) Familial systemic paramyloidosis with lattice dystrophy of the cornea, progressive cranial neuropathy, skin changes and various internal symptoms. A previously unrecognized heritable syndrome. *Ann. Clin. Res.*, **1**, 314–324.
86. Haltia, M., Prelli, F., Ghiso, J., Kiuru, S., Somer, H., Palo, J. and Frangione, B. (1990) Amyloid protein in familial amyloidosis (Finnish type) is homologous to gelsolin, an actin-binding protein. *Biochem. Biophys. Res. Commun.*, **167**, 927–932.
87. de la Chapelle, A., Tolvanen, R., Boysen, G., Santavy, J., Bleeker-Wagemakers, L., Maury, C.P.J. and Kere, J. (1992) Gelsolin-derived familial amyloidosis caused by asparagine or tyrosine substitution for aspartic acid at residue 187. *Nature Genet.*, **2**, 157–160.
88. Kwiatkowski, D.J., Westbrook, C.A., Bruns, G.A.P. and Morton, C.C. (1988) Localization of gelsolin proximal to ABL on chromosome 9. *Am. J. Hum. Genet.*, **42**, 565–572.
89. Paunio, T., Kangas, H., Kalkkinen, N., Haltia, M., Palo, J. and Peltonen, L. (1994) Toward understanding the pathogenic mechanisms in gelsolin-related amyloidosis: *in vitro* expression reveals an abnormal gelsolin fragment. *Hum. Mol. Genet.*, **3**, 2223–2229.
90. Paunio, T., Kiuru, S., Hongell, V., Mustonen, E., Syvänen, A.-C., Bengström, M., Palo, J. and Peltonen, L. (1992) Solid-phase minisequencing test reveals Asp187→Asn (G654→A) mutation of gelsolin in all affected individuals with Finnish type of familial amyloidosis. *Genomics*, **13**, 237–239.
91. Paunio, T., Kangas, H., Heinonen, O., Buc-Caron, M.-H., Robert, J.J., Kaasinen, S., Julkunen, I., Mallet, J. and Peltonen, L. (1998) Cells of the neuronal lineage play a major role in the generation of amyloid precursor fragments in gelsolin-related amyloidosis. *J. Biol. Chem.*, **273**, 16319–16324.
92. Hallman, N., Norio, R. and Kouvalainen, K. (1967) Main features in the congenital nephrotic syndrome. *Acta Paediat. Scand.*, **172** (suppl.), 75–77.
93. Holmberg, C., Jalanko, H., Tryggvason, K. and Rapola, J. (1999) Congenital nephrotic syndrome. In Barratt, T., Avner, E. and Harmon, W. (eds), *Pediatric Nephrology*, 4th Edn. Lippincot William and Wilkins, Baltimore, MD, pp. 765–778.
94. Hästbacka, J., Kerrebrock, A., Makkala, K., Clines, G., Lovett, M., Kaitila, I., de la Chapelle, A. and Lander, E.S. (1999) Identification of the Finnish founder mutation for diastrophic dysplasia (DTD). *Eur. J. Hum. Genet.*, in press.
95. Aittomäki, K. (1994) The genetics of XX gonadal dysgenesis. *Am. J. Hum. Genet.*, **54**, 844–851.
96. Simoni, M., Gromoll, J. and Nieschlag, E. (1997) The follicle-stimulating hormone receptor: biochemistry, molecular biology, physiology, and pathophysiology. *Endocrine Rev.*, **18**, 739–773.
97. Mishari, M., Beau, I., Meduri, G., Bouvattier, C., Atger, M., Loosfelt, H., Ghinea, N., Hai, M.V., Bougneres, P.F. and Milgrom, E. (1998) Gonadotropin receptors and the control of gonadal steroidogenesis: physiology and pathology. *Baillieres Clin. Endocrinol. Metab.*, **12**, 35–66.
98. Tapanainen, J.S., Aittomäki, K., Min, J., Vaskivuo, T. and Huhtaniemi, I.T. (1997) Men homozygous for an inactivating mutation of the follicle-stimulating hormone (FSH) receptor gene present variable suppression of spermatogenesis and fertility. *Nature Genet.*, **15**, 205–206.
99. Gibson, T.J., Ramu, C., Gemund, C. and Aasland, R. (1998) AIRE-1, involved in the polyglandular autoimmune syndrome APECED, contains the SAND domain and is a probable DNA-binding transcription factor. *Trends Biochem. Sci.*, **23**, 242–244.
100. Kere, J., Lohi, H. and Höglund, P. (1999) Genetic disorders of membrane transport—III. Congenital chloride diarrhea. *Am. J. Physiol. Gastrointest. Liver Physiol.*, **39G**, G7–G13.
101. Huopaniemi, L., Rantala, A., Forsius, H., Somer, M., de la Chapelle, A. and Alitalo, T. (1999) Three widespread founder mutations contribute to high incidence of X-linked juvenile retinoschisis in Finland. *Eur. J. Hum. Genet.*, **7**, 368–376.
102. Rinderle, C., Christensen, H.M., Schweiger, S., Lehrach, H. and Yaspo, M.L. (1999) AIRE encodes a nuclear protein co-localizing with cytoskeletal filaments: altered sub-cellular distribution of mutants lacking the PHD zinc finger. *Hum. Mol. Genet.*, **8**, 277–290.
103. Visapää, I., Salonen, R., Varilo, T., Paavola, P. and Peltonen, L. (1999) *Am. J. Hum. Genet.*, in press.