Chapter 3

Literature Review

The human genetics study started with the comprehensive theory of inheritance given by In Aristotle in 384-322 B.C. He believed in a qualitatively different contribution by the male and the female principles to procreation. “The male gives the impulse to movement whereas the female contribute the matter. When the male impact is strong, a son is born who,at the same time,is more like his father, when female,a daughter,resembling the mother”. Though according to him a son is more like a father and daughter is more like a mother [9].The medical literature of eighteenth and early nineteenth centuries contains reports showing that they were capable of clear observation and were able to correctly recognize some phenomenon relating to inheritance of disease. For example Maupertuis , a mathematician described an autosomal dominant polydactyly in 4 generations of a family and demonstrated that traits could be equally transmitted by father or by mother [10].In 1865 F. Galton in his book “Herediary Talent and character” discussed that single observation of the characters can be misleading and only statistical approaches can be applied to study inheritance [11].

Gregor Mendel from the Czechoslovakia sovereign state in Central Europe spend his lifetime was the first scientist to deduce clear and rational laws which could explain the process of inheritance in humans. It is convenient to trail his logic from characterizing a single gene followed by moving on to the complications introduced by multiple genes. The results of Mendel's experiments are still the backbone of pedigree analysis [12].

For the first time in year 1902,Mendel’s gene concept was applied to human characters,and Mendels paradigm was introduced by A Garrod in his paper “The Incidence of Alkaptonuria:A study in chemical Individuality”. The first inborn errors described early in the 20th century came mainly from the studies on the rare disorder Alkaptonuria. This was a relatively benign disorder but often diagnosed in infancy because of brown discoloration in urine .Initial physical and chemical test where used to analyze the disease process in the family[13] .

3.1 Gossage styled pedigree

In the year 1907 Gossage reported a family with heterochromia of the iris(Figure 3.1) ,which tends to appear when one eye, always the left, was grayish blue in color, with chestnut-brown patches.

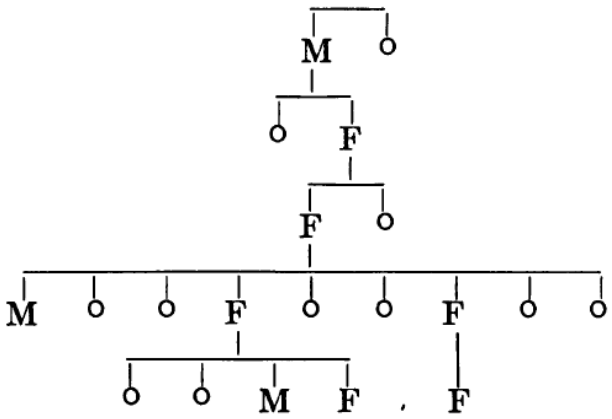

Figure 3.1 Gossage styled pedigree.

This was a 281 member pedigree.Of all the offsprings, eight were affected and twenty-two were disease free[14].

3.2 Batten and Mayou styled pedigree

The Batten described symmetrical changes in the macule in two members of a family with cerebral degeneration starting about the age of 6 were described in the year 1903. In 1904 M. S. Mayou described cerebral degeneration with symmetrical changes in the macula in three members of a family starting about the age of 7.

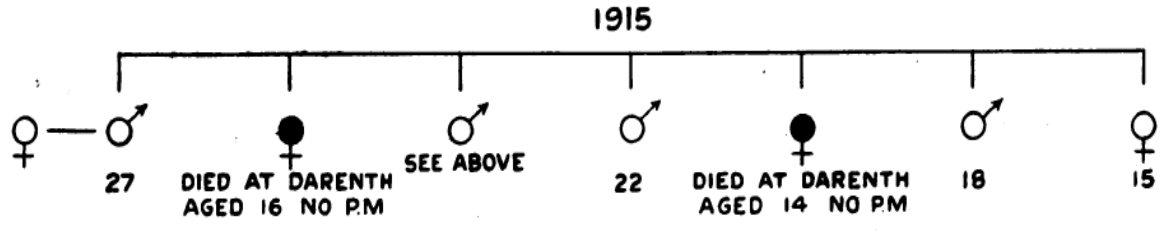

Figure 3.2 Batten and Mayou styled pedigree.

Utimately in the year 1915 together gave a pedigree image(Figure 3.2) for Cerebral Degeneration with Macular Changes[15].

3.3 Royal British Pedigree

The pedigree chart of Queen Victoria of England(Figure 3.3) illustrates inheritance of hemophilia A. Hemophilia was the royal disease of 18th century, having a X-linked mode of inheritance [16]. Since the characteristic symptom of haemophilia was the non-functioning of the joints, it was many a times confused with tuberculous, rheumatic, and other types of arthritis.. Queen Victoria herself was a carrier due to a chance mutation. Her children married other royalty and passed the trait throughout the royal families of Europe.In Victoria's family two of her daughters proved to be carriers, transmitting the disorder to three of her grandsons and six of her great-grandsons. The condition was not known among any of the Queen's antecedants, so it was supposed that a mutation occurred at spermatogenesis in her father, Edward, Duke of Kent.The mutation was perhaps more probable because of the fact that he was in his fifties when she was conceived [17].

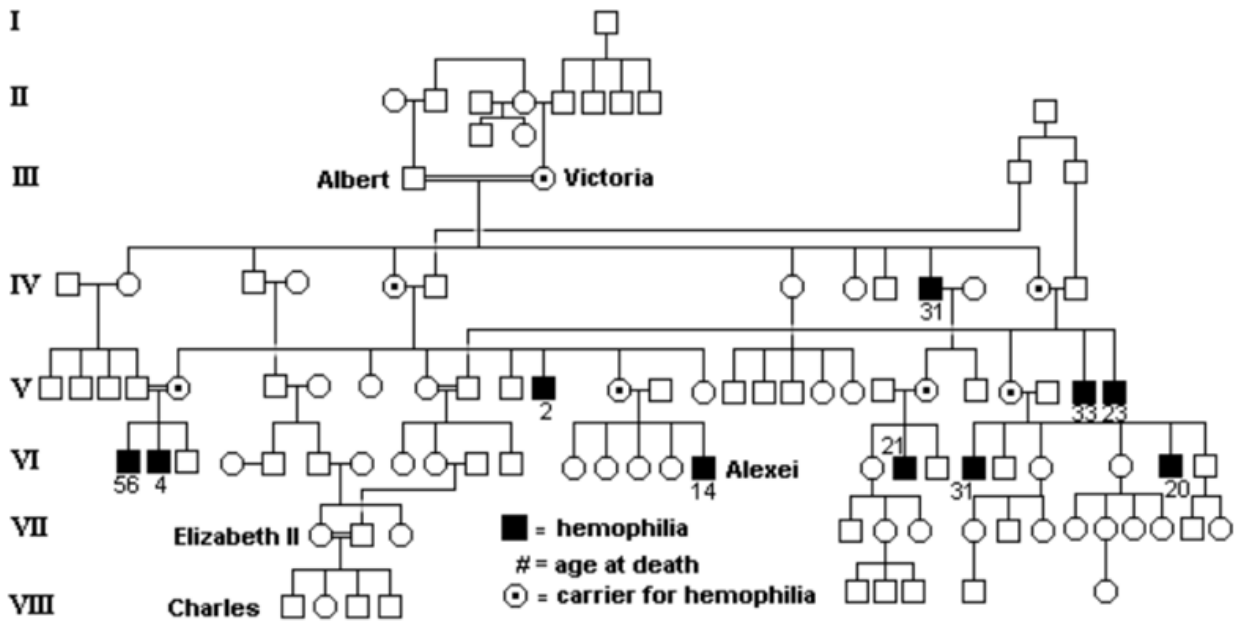

Figure 3.3 Royal British Pedigree.

3.4 Reconnection of Nuclear LHON-Affected Families into Extended Pedigrees

Leber hereditary optic neuropathy (LHON) was the first disease to be linked with an mtDNA point mutation [18] and is characterized by a maternally inherited loss of central vision, which occurs most frequently in males. As the mitochondrial genome is small but highly mutable compared to nuclear DNA, probably because mitochondrial DNA replication is more error-prone and the number of replications is much higher. Three mtDNA point mutations affecting the ND (NADH dehydrogenase) subunits of the respiratory enzyme complex I (11778/ND4, 3460/ND1, and 14484/ND6) are commonly found worldwide and are well established as pathogenic. In figure the reconstructed pedigree generated from families BSL07 and LB10 is shown. The BSL07 family originates from Brazil and is a nuclear subset of the large Brazilian SOA-BR pedigree, which has been completely reconstructed and investigated from a genetic, epidemiological, and clinical point of view. The initial SOA-BR family (Figure 3.4) has approximately ∼360 family members, 128 of whom were maternally related. Of the latter, 35 are affected (23 are alive, 2 of whom are recent cases and became affected in the past 2 years). All 122 individuals investigated are homoplasmic for the 11778/ND4 mutation. The woman—daughter of the most recent common female ancestor of the entire pedigree, who migrated to Brazil and gave rise to the Brazilian branch—was born in 1861 in the Veneto region of northern Italy, the same geographical area from which the LB10 family originates [19].

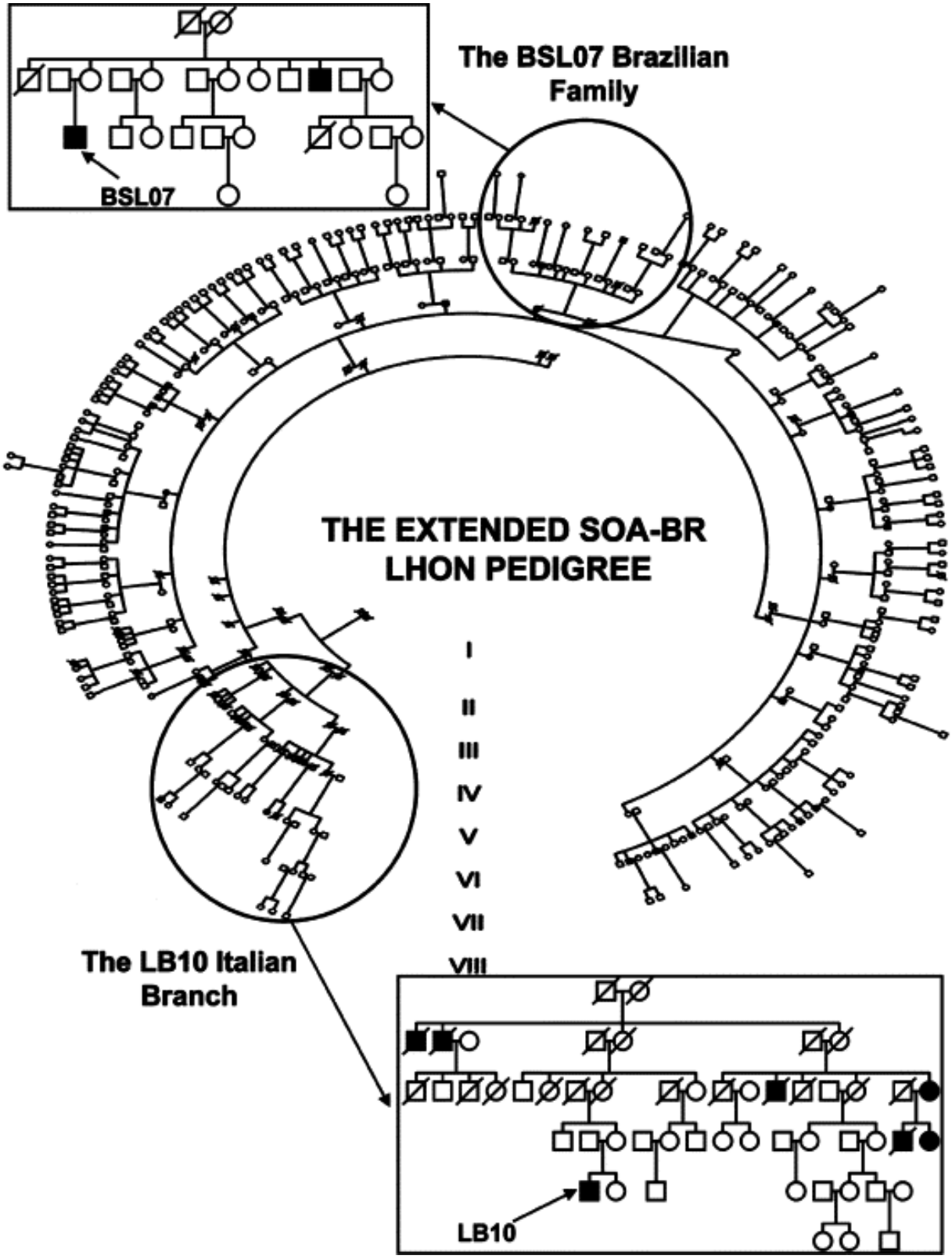

Figure 3.4 LHON pedigree of Brazilian and Italian.

3.5 Genetic Epidemiology

Almost 60 years after Mendel's theory a new field of science, **Genetic Epidemiology** emerged as a hybrid of genetics, biostatistics, epidemiology and molecular biology, which played a major tool in establishing whether a phenotype has a genetic component attached to it and also the measure of relative size of that genetic effect in relation to environmental effects. Morton and Chung defined genetic epidemiology as "a science that deals with the etiology, distribution, and control of disease in groups of relatives, and with inherited causes of disease in populations" [20]. Traditionally, the study of the role of genetics in disease progresses through the pedigree study can provide answer to many questions like the genetic component to the disease, and the relative contributions of genes and environment, the pattern of inheritance of the disease, part of chromosome associated with the disease or even location of the gene.

3.5.1 A Family Aggregation Study: The Influence of Family History and Other Risk Factors

Familial aggregation for a disease is measured by the relative recurrence risk (RRR) or familial risk ratios (FRRs). Examination of relative recurrence risk values for various classes of relatives can potentially suggest a polygenic background and epistasis [21].

The familial aggregation is the occurrence of more cases of a given disorder in close relatives of a person with the disorder than in control families. Although this condition has been shown for many (if not all diseases) but time and again it has become problematic to define whether such aggregation is due to shared genes or shared environmental factors. The researchers have examined the effect of familial aggregation of environmental risk factors on familial aggregation of disease when there is no genetic susceptibility [22].

A basic method for determining the pattern of inheritance of any trait is to look at its occurrence in several individuals within a family, spanning as many generations as possible. For a disease trait, a doctor has to examine existing family members to determine who is affected and who is not. The same information may be difficult to obtain about more distant relatives, and is often incomplete. The initial stage of model building includes choice of traits describing the biological function to be studied. The mathematical model is formulated and tested using pedigree samples. It is supposed that implicit details of the trait's mode of inheritance through co-variance and co-segregation of the traits are characteristics among the

members of the pedigree. The study describes the family aggregation and demographic characteristics of a unique group of individual who were born into families with a known history of disorder. The studies provide information about the extent to which disorder runs in families. This type of studies has been carried out on numerous disorders and impairments [23, 24]. Family history is a major risk issue for patients with an affected first-degree relative who have increased risk of colorectal cancer (CRC) [25]. Most often, such studies use a retrospective approach to collect family history information. Thus, after identifying a proband (the affected individual), parents or individual family members are asked to complete questionnaires on the performance of family members. In case of Specific language impairment (SLI), the results of the studies have suggested a robust effect ; approximately 40% to 60% of families have a pro-band with impairments in at least one or the other immediate family member [26, 27, 28].

The results of observational studies on siblings, parent-offspring concordance, twins, adopters and even migrants suggest a genetic component in the etiology of a disease or trait. Familial aggregation of a trait is a necessary but not sufficient condition to infer the importance of genetic susceptibility, because environmental factors also aggregate in families, leading to family clustering and excess familial risk. Similar environment may be the reason for familial aggregation. With rising divorce rates, study of recurrence risk in half-siblings is another powerful method to test for parent-specific events. In an application of this method, multiple sclerosis appeared to have a genetic basis transmitted more from mothers than fathers [29].

3.5.2 Segregation studies

Segregation analysis is a prerequisite for linkage analyses. Segregation analysis reveals Mendelian inheritance patterns (autosomal or sex-linked and recessive or dominant); nonclassical inheritance (mitochondrial diseases, genomic imprinting, parent of origin effect, genetic anticipation etc); or non-Mendelian inheritance (no pattern) .Factors interfering with genotype-phenotype correlation such as incomplete penetrance, variable expressivity, confounding by other genes (allelic or locus heterogeneity, multigene inheritance, epistasis, modifier genes, sex influence, parental effect) or environmental factors, and nonclassic genetic phenomena (imprinting - functional differences exist between the paternal allele and the maternal allele, which is a result of process of epigenetic regulation, mitochondrial

inheritance) complicate the segregation analysis of complex diseases for which no inheritance pattern is obvious despite familial aggregation. **Complex segregation analyses** are based on more elaborate mathematical methods of genetic transmission and liability [30].

The segregation studies include studying the pattern of inheritance of the disease condition. The scientist have applied various physical and chemical test to analyze the disease process in the family .The Maroteaux in the year 1971 reported two families with metachondromatosis disorder and suggested autosomal dominant inheritance on the basis of 1 family with 5 affected persons

This study was carried out by analyzing the skeletal radiologic features of both multiple exostoses and Ollier disease which are benign cartilaginous tumors, that develop in close proximity to growth plate cartilage [31]. In the year 2000, Kramer et al. studied 12 affected and 13 unaffected members of a 4-generation family with autosomal dominant congenital cataract. Affected family members had congenital nuclear and suture cataracts of variable severity, with some family members undergoing surgery in childhood and others never requiring surgery [32].

The first disease-causing mutations in Rett Syndrome, was found by Amir et al. in the year 1999.The abnormal epigenetic regulation was underlyed as the mechanism for causing pathogenesis in RTT .It is a progressive neurodevelopmental disorder and one of the most common causes of mental retardation in females, with an incidence of 1 in 10,000-15,000 . Previous exclusion mapping studies using RTT families mapped the mutation to Xq28 locus . Using a **systematic gene screening approach**, it was identified that the mutations in MECP2 gene encoding X-linked methyl-CpG-binding protein 2 (MeCP2) is the major cause for some cases of RTT. In 5 of 21 sporadic patients, 3 de novo missense mutations in the region encoding the highly conserved methyl-binding domain (MBD) as well as a de novo frameshift and a de novo nonsense mutation, both of which disrupt the transcription repression domain (TRD) was found . In two affected half-sisters of a RTT family, segregation of an additional missense mutation was seen which was not detected in their obligate carrier mother. This suggests that the mother is a germline mosaic for this mutation.Only a few disorders have this inheritance pattern, like X-linked hypophosphatemic rickets. The degree of disruption in the functioning of the enzyme was similar in case of males and females, although the severity of bone disease was much less severe in females. There were no occurrences of male-to-male transmission of either bone disease or

hypophosphatemia, and all daughters of hypophosphatemic males were themselves hypophosphatemic, suggesting X-linked dominant inheritance [33].

A higher concordance rate for brother pairs than in father-son pairs suggests an X-linked recessive genetic background (as has been observed X-linked prostate cancer [34]). Same sex pairs of affected relatives suggest X-linked recessive (males) or X-linked dominant (females; due to intrauterine loss of males) inheritance. The genes within pseudoautosomal regions (PAR) of the X and Y chromosomes have a unique segregation pattern that the disease can occur in either sex but affected sibs tend to be same sex. The mutant allele would consistently segregate, during male meiosis, with sexual phenotype.

The X-inactivation (lyonization) blurs the distinction between dominant and recessive X-linked conditions. Inactivation takes place early in embryonic life, and once a cell has chosen which X to inactivate, that choice is transmitted clonally to all its daughter cells. A female heterozygous for an X-linked condition (dominant or recessive), is a mosaic . In these heterozygotes infrequent X-linked recessive conditions are seen: these women may be quite severely affected because by bad luck most cells in some critical tissue have inactivated the normal X.

A man could possess the mutant allele on either his X or Y chromosome. If it is on the X chromosome, then only his daughters would inherit the allele, whereas if it resided on the Y chromosome, then only his sons would inherit the allele [35]. Although a few Y-linked characters have been described, no Y-linked diseases are known, apart from disorders of male sexual function. Conceivably such a disease may exist undiscovered, but this is unlikely for two reasons. First, the pedigree pattern would be strikingly noticeable, especially in societies that trace family through the male line, yet they have not been noted (claims for ‘porcupine men’). Second, the Y-chromosome cannot carry any genes whose function is important for health, because females are perfectly normal without any Y-linked genes. Thus any Y-linked genes must code either for non-essential characters or for male-specific functions, and defects are unlikely to cause diseases apart from defects of male sexual function. In addition to the alterations in genes carried on the nuclear chromosomes, mitochondrial mutations are also a significant cause of human genetic disease.

Mitochondrially-encoded diseases have two unusual features, maternally inheritance [36] and frequent heteroplasmy [37] refers to mosaicism, usually within a single cell, for mitochondrial DNA variants. This type of inheritance, also known as maternal inheritance, applies to genes in mitochondrial DNA. Because only egg cells contribute mitochondria to the developing embryo, only mothers can pass on mitochondrial conditions to their children. Thus a Mitochondrially inherited condition can affect both sexes, but is passed on only by affected mothers gives a recognizable pedigree pattern.

In 2002 Sedlmeyer investigated the varied inheritance pattern in genes involved in constitutional delay of growth and maturation (CD) in 41 families of Boston. Several different inheritance patterns were observed in the study of these population suggests that multiple genes are likely modulate the timing of puberty in humans. In addition, some pedigrees could theoretically be attributed to more than one mode of inheritance. For example, some pedigrees classified as autosomal recessive could actually represent autosomal dominant inheritance with incomplete penetrance in the parent carrying the causal genetic variant [389. A particularly important case of reduced penetrance is seen with late-onset diseases. Genetic conditions are, of course, not necessarily congenital. In such cases the genotype is fixed at conception, but the phenotype may not manifest until adult life. The penetrance is age-related like in Huntington disease. Delayed onset might be caused by slow accumulation of a noxious substance, by slow tissue death or by inability to repair some form of environmental damage.

In some case the mode of inheritance cannot be ascertained. Zachara et al. in the year 1993 studied the family history of 105 patients with dilated cardiomyopathy .The study gave new insights into the aetiology, the pathogenesis, and the natural history of patients with dilated cardiomyopathy. The mode of inheritance was not be ascertained as cardiomyopathy is not consistent with being entirely genic but with being polygenic accounting only in part for disease susceptibility, the remainder being related to environmental factors such was seen in autoimmune conditions[31]. The analysis of family with type 2 neurofibromatosis having linkage with neurofilament heavy chain (NEFH) locus was carried out. The two affected offspring inherit no NEFH allele from their mother. The likely cause of both the disease and the anomalous inheritance pattern is a deletion encompassing both the NEFH gene and NF2 genes [40] .

3.5.3 Strategies in identifying disease genes

In 2 families by analysis skeletal radiologic features of both multiple exostoses and Ollier disease, the disorder as metachondromatosis was reported and autosomal dominant inheritance was observed on the basis of 1 family with 5 affected persons [31].In the 1980s, advances in 'reverse genetics' lead to an increase in the identification of genetic factor behind numerous disease. With the advent of PCR for linkage studies and mutation screening, it all became much easier. With the success of human genome projects, a vast range of resources - maps, clones, sequences, expression data and phenotypic data was made available for research purpose. Historically, the first disease genes were identified by pure position-independent methods, simply because no relevant mapping information existed and the techniques were not available to generate it.

Later-on the protein, or a peptide derived from it, is conjugated to a powerful immunogenic hapten was used to identify the gene .This approach was applied for gene identififcation of phenylketonuria [41]. Phenylketonuria was known to be caused by a lack of the enzyme phenylalanine hydroxylase (PAH).

The cystic fibrosis locus was identified by using a 'rare-cutter cosmid library.' They found a genomic region with the characteristics of an HTF island in high linkage disequilibrium with CF. The fact that the sequence was conserved throughout mammalian evolution strengthens the view that this is the CF gene. HTF islands, standing for HpaII tiny fragments, have a sequence length of between 500 and 1000 bp and often include the first exons as well as upstream sequences 5-prime to coding genes [42].

After this most of the researchers shifted to use the repeat expansion detection method of Schalling *et al*. (1993) for detection of expanded repeats in unfractionated genomic DNA [43] of affected patients, and methods was developed for cloning any expanded repeats detected . This approach was used in to identify a novel repeat expansion that causes a form of spinocerebellar ataxia (SCA8) [44].

In recent times the first successful gene identifications based only on positional information, published in 1986, left a mark in the field of genetics. One after another, genes for important disorders such as Duchenne muscular dystrophy, cystic fibrosis, Huntington disease, adult

polycystic kidney disease, colorectal cancer, breast cancer, etc. were isolated. However, positional cloning can be desperately hard work, and by 1995 only about 50 inherited disease genes had been identified by this approach [45].

3.5.4 Linkage Studies

Linkage analysis in man has long been restricted to simple Mendelian traits and nuclear families. In 1934 Fisher used simultaneous estimation methods in the evaluation of linkage [46].Soon after him in the year 1940, Finney carried out the work on detection of genetic characters and its importance in studying human heredity. According to him linkage relations between heritable abnormal conditions and common markers genes is a prerequisite for preventing the spread of harmful abnormalities [47]. Seven years after Finney, Haldane along with his co-worker Smith showed the new estimate of the linkage between the genes for haemophilia and colour-blindness in man [48, 49].Morton in 1955 used the sequential probability ratio test to observe the least number samples consistent with a given risk [50].

In 1993, linkage between homosexuality and chromosomal region Xq28 based on molecular approaches was reported. But still much is to be known about the genetics of human homosexuality [51].

3.6 Genetic association studies

Genetic association studies, are designed in order to select polymorphic markers within candidate genes or the genic regions and within that region the extent of allelic association with disease at those markers are measured within a case-control or family-based design. Whereas linkage studies are family-based studies that measure the co-segregation of trait loci with genetic markers within each family under study. Genome-wide linkage studies (GWLS), by extension, use a huge set of hundreds to thousands of DNA markers that are evenly spaced across the genome to detect broad regions likely to cause disease susceptibility loci, based on the pattern of inheritance seen within-family correlations between marker alleles and disease [52, 53, 54].

Genetic association studies, were designed in order to select polymorphic markers within candidate genes or the genic regions and within that region the extent of allelic association with disease at those markers were measured within a case-control or family-based design. Also Classical Hodgkin lymphoma (cHL), malignancy of B-cell origin in which the neoplastic cells was described using this technique. Mapping the chromosomal translocation breakpoints in Metaphase FISH of lymphoblastoid cells localized breakpoints to a segment on chromosome 3 containing two genes and to a region on chromosome 2 lacking apparent genes.Several members in the family developed the disorder due to reciprocal translocation between chromosomes 2 and 3. This translocation disrupted the functioning of *KLHDC8B*, an uncharacterized gene from a region (3p21.31) previously implicated in lymphoma and related malignancies, resulting in its loss of expression [55].

3.7 Heritability

Heritability is the proportion of phenotypic variation in a population that is due to genetic variation between individuals. A very few studies have focused on the amount of familial aggregation and heritability of diseases in a population. For this all risk factors showing positive familial correlation are studied. The analysis is carried by multivariate model which includes the each individual family data for estimation of familial correlation and heritability. One such done on factors leading to cardiovascular disease showed correlations were higher among genetically close relatives such as brother-sisters or parent-offspring and were lower among spouses [56,57,58].

3.8 Genetic epidemiology of complex diseases

Human sexual orientation is a complex trait, influenced by quite a few genes, environmental and socio-cultural factors. These elements interact and produce a typical pattern of sexual orientation towards the opposite sex. Some exceptions exist, like bisexuality and homosexuality, which seem to be more frequent in males than females. Conventional methods for the genetic study of multifactorial characteristics of the behaviour consist in detecting the existence of familial aggregation. In order to identify the importance of genetic and environmental factors in this aggregation, the concordance of the trait for monozygotic and dizygotic twins and for adopted sibs, reared together and apart, were compared. These

types of studies have shown that familial aggregation is stronger for male than for female homosexuality [51].

A statistical technique to maximize the probability of localizing disease gene was carried out using population-based association .For the gene-mapping studies, either specific alleles, genotypes, or haplotypes was collected from unrelated cases and controls. The data was then used for testing the associations between the genes and the complex traits [60].

According to the fact sheet Fact Sheet No 209, January 1999 of WHO information, about 5% of progenies are born with a congenital or hereditary disorder and almost 40% of adults are genetically predisposed to common diseases during their life-time. Even in developed countries, quarter of deaths under the age of one and 23% between one and four suffer from congenital and genetic disorders. In recent years, genetic studies have shifted from single locus disorders (e.g. Phenylketonuria disease) to polygenic disorders (e.g. hypertension) that result from the interactions between inherited gene variants and environmental factors, including chemical, physical, biological, social, infectious, behavioral or nutritional factors.

In the era of a known human genome sequence, methods have been instrumental in identifying the impact of genes, the environment and their interactions to better understanding disease processes. Traditionally, the role of genetics in a disease progresses passes through four components. The main component is the knowledge of pattern of inheritance of the disease .Other components includes the research of genetic component to the disease, and relative contributions of genes and environment in the occurrence of the disease. The approaches to identifying genes and genomic regions associated with human disease can be grouped into two categories: linkage analysis and genetic association analysis. The information about the location of the gene on the chromosome; the allele associated with the defect. These approaches have been successful in identifying monogenic disorders and locating the genes responsible.

By characterizing genetic variation among individuals and populations, we may gain a better understanding of differential susceptibility to disease, differential response to pharmacological agents, human evolutionary history, and the complex interaction of genetic and environmental factors in producing phenotypes.

The advantage of analyzing family pedigrees is that disease inheritance can be compared to patterns of linkage over large genomic regions in an effort to map genetic mutations. On the other hand, it is difficult to collect affected families of sufficient size or number to effectively apply these methods. Finally, traditional linkage studies might have limited power for identifying the moderate gene effects postulated to contribute to complex diseases Often two experimental strategies are combined together to investigate genetic variants in human disease: linkage analysis and association studies. The former, linkage analysis, seeks to define a physical relationship between two or more genetic markers, identifying the locus of a disease gene, whereas association analysis correlate a sequence variant with a well-defined phenotype. Linkage analysis compares inheritance patterns of predefined genetic markers residing on the same chromosome to a disease outcome. The approach pinpoints a region of a chromosome based upon a non-random pattern of co-inheritance of markers, although usually it does not uncover the specific genetic variation responsible for disease outcome. For tracing segregation patterns of genetic markers multi-generation family pedigrees are required .Linkage studies have been employed to map hundreds of highly penetrant disease loci[29].An association studies are used to test whether a SNP, repeat region or microsatellite is enriched in patients with disease compared to suitable controls.