

DNA Profiling:

Genetic polymorphism in human species and other living systems is the hallmark of variations. For example hyper variable loci show extraordinarily individualistic for a given individual. Hyper variable loci are made up of variable number of identical sequences are arranged in tandem fashion. If such hyper variable (HPV) or variable number of tandem repeat (VNTR) regions are cut at their flanking region and analyzed on a n Agarose or acrylamide gel variation in individual lengths of such polymorphic DNA is visible. Such distinct number of bands differs from one individual to the other with the exception of monozygotic twins. The composition and variability of VNTR loci is an expression high degree of polymorphism.

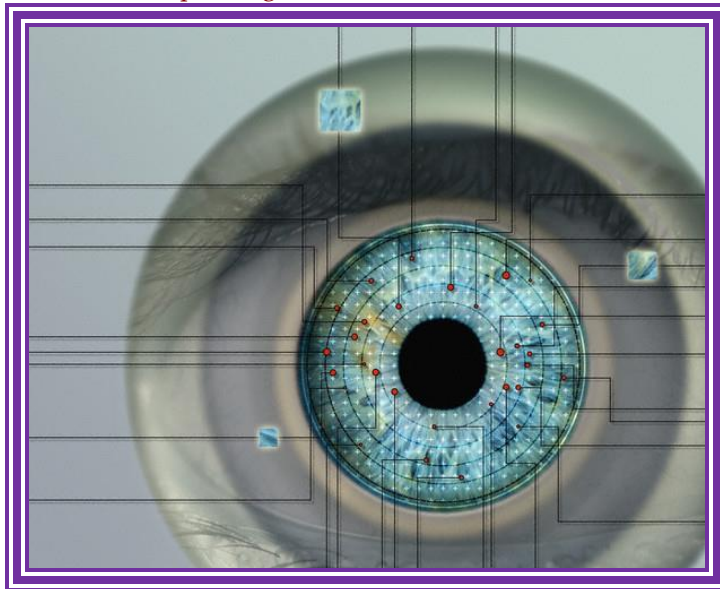
The first HPV probe used for individual identification consisted of 33bp found in introns of Myoglobin gene. It actually had 10-15 bp long core sequence. When two such clones were used as probes 30-40 distinct bands could be discerned on a southern blot.



Physical finger printing, an age old method, still being used



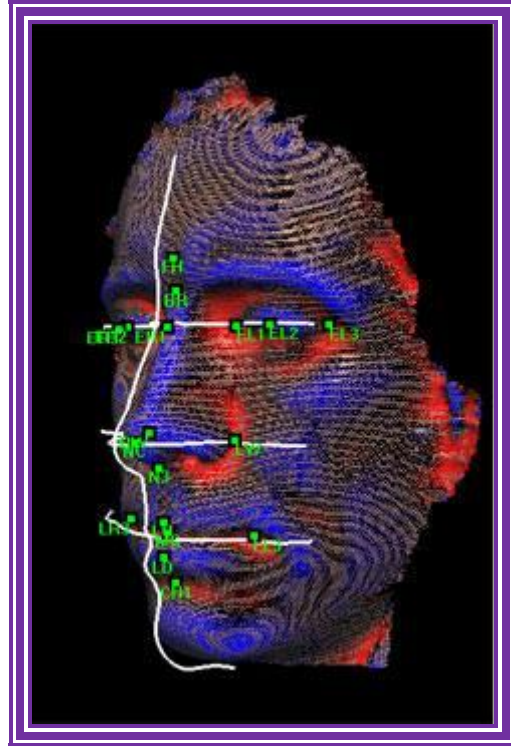
Imprinting machine; **Retinal scanner**



Retinal scanning



Iris scanning



Face imaging

Many of the countries use biometric methods for personal identifications. They can use retinal scanning (popularly done in USA when you enter any of the airports at Immigration desk). Many of the biometric techniques often use facial characteristics.

DNA profiling:

Alec Jeffrey and his colleagues (1985) developed this technique, which has far reaching impact in molecular biology. This technique is often referred to as "Digital Finger Printing". They found a sequence of nucleotides in human hemoglobin DNA was repeated several times. Furthermore, such sequences are also found elsewhere in the genome. Such sequence elements, which are tandemly repeated in different loci, are called mini-satellites, which can be isolated by ultra centrifugation techniques. The sequences found were found to be made up of 2, 3, 4, 6, 7, 8 and 9 and ten or more nucleotide sequences. And each of the elements is tandemly repeated and the number of repeats for each species can vary from 10 to 100 or more and the number of repeats per locus varies from 5 to 50 or more. This provides variability and this variability of the said elements is species specific and fixed.

2bp: (GT / AC) n, (AT / TA) n, (CT / GA) n,
3bp: (CTG) n, (CAC) n, (GAA) n, (TCC) n,
4bp: (ATCC) n, (GATA) n, (GACA) n, (GGAT) n,
5bp: (AGAAG) n,
9bp: (GGG CAG GNG) n.

Perfect micro satellites (GT) n:

GT,
GT GT,
GT GT GT,
GT GT GT GT,
GT GT GT GT GT,
GT GT GT GT GT GT,

Interrupted MS:

GT GT GT GA GT GT,

Composite MS:

GT GT GT GT CT CT CT CT,

Cryptic MS:

GA GT GT CTT CTT GTC TG TGTTT G.

Y linked MS:

The denoted 'n' can be any number that can be from one to thousand tandem repeats. A given locus consists of a particular sequence and specific number of repeats, this is a constant feature of the locus and the number of repeats may vary in its homologous locus. Homologous chromosomes may contain many such loci having different characteristic features. Distinct loci with different sequences and with different number of repeats have been identified.

A 33bp probe was used for identification of introns in Myoglobin gene in human species. Using this probe people identified many

clones from the human genomic cosmid library. Such sequences were found to be 10-15 bp long. One of the repeats such as CA / GT has been found to have 77 alleles in a population and the number of repeats has been found to range from 14 to 500. Such CA/GT has been found to occur in about 100,000 blocks in human genome. Even islands of GC/CG sequences are also found mostly in the regulator regions. Any such segments can be easily amplified by PCR.

For example, PYNH24 probe, developed from NIH (?), when used on 16 unrelated people detected 19 alleles and no two individuals contained the same pattern

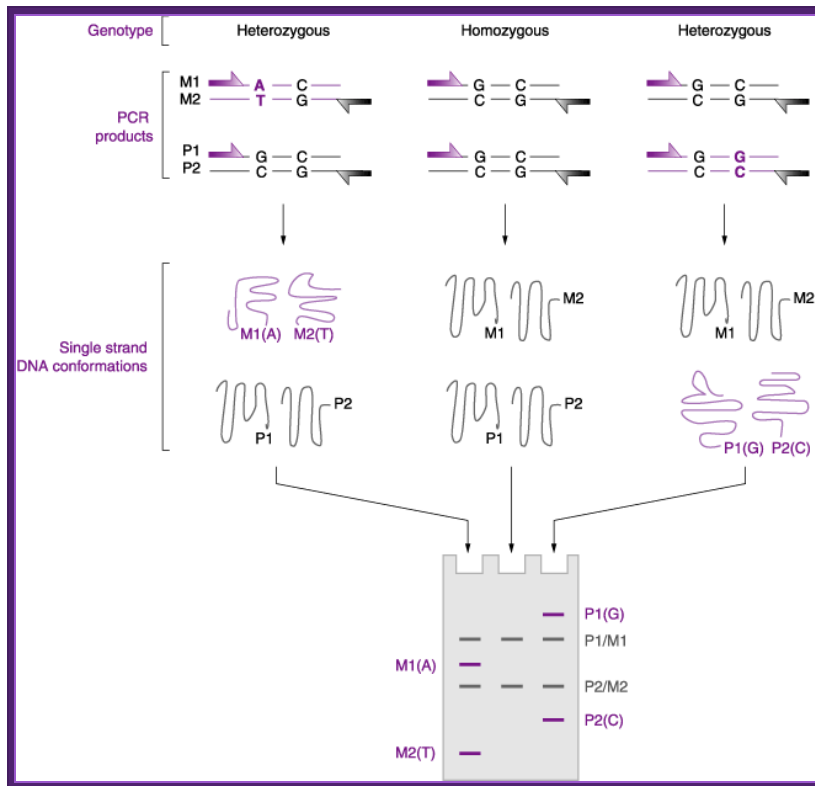
Such repeats originate by unequal crossing over, which can spread this sequence all over, the chromosomal length. Some of the sequences may increase in their number and some sequences may shrink in their numbers. Nearly hundred of distinct min-satellite has been identified and they are distinctly different.

Let us look at the inheritance pattern of heterozygous alleles of these repeats. For example, a single locus contains two alleles, say male has 6bp and 9bp long and female consists of 5bp and 7bp long.

Multilocus fingerprints contain tandem repeat arrays and they are found in various region of the genome. When such Multilocus DNA the result appears as bar codes used in supermarket commodities used for identification of the product and its price.

Cross between two allelic genotypes:

Male/female	6n	7n
5n	5n/6n	5n/7n
9n	9n/6n	9n/7n



DNA profiling

If the number of alleles is more than two, which is actually the case, the identity of alleles of a single individual, becomes easier. For forensic analysis a large number of markers have been developed. They are short tandem repeats (STRs), which are derived from micro satellite DNA. They are made up of two to nine base pair long sequences and found in abundance in human genome, which is also true for other species. Using less than 1ng of DNA one can amplify it by PCR overnight.

Probes for Micro satellite and Minisatellite for single or multilocus have been developed.

The STR alleles consist of 100 to 350 ntds. Using multiplex PCR one can amplify many STR loci simultaneously. In USA 13 STR loci

are employed for typing all felonies using National Data Bank called CODIS.

The 13 loci probes are- CSF1PO, D3S1358, D5S818, D7S820, FGA, THO1, TPOX, VWA, D8S1179, D13S317, D16S539, D18551, D21S11. Perkin Elmer Company has developed three multiflour multiplex STR kits and Promega people have developed 8 loci STR kits. Each kit provides fluorescent dye-labels, a multiplex primer sets and reagents for PCR amplification.

Amplified DNA can be analyzed on PGA (poly acrylamide gel) or capillary electrophoresis; the later is good, faster and automated. Capillary tubes with the gel inside have high surface area to volume ratio, thus heat is dissipated. The bands are separated at high field strength and the results are computer analyzed and the data is stored. The capillary gel consists of Dimethyl polyacrylamide (POP4), in 100mM TAPs, 8M urea and 5% pyrrolidine and chromogenic dyes such as 5-FAM-blue, JOE-green, ROX- red and NEO-Yellow. The enzyme used is Ampli-Taq-Gold.

Highest individualism has been obtained by using (CAC) 5 and its complement (GTC) 5.

Some of the sequences effectively used are-

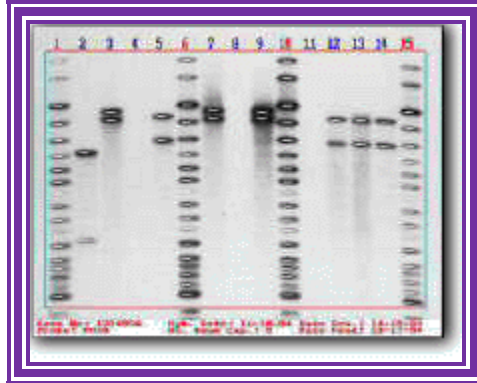
- (CA) n,
- (CT) n,
- (GAA) n,
- (TCC) n,
- (GGAT) n,
- (GATA) n,
- (GACA) n,

Different Individuals:

(CAC) 5

-			
=	-		
=	-	=	
-	=		
=	-		
=	-		
-	=		
-	=		
-	=		
=	-		
=	-		
=	-		
-	=		
=	-		
-	=		

Note that none of the individuals have copies of the same sequence pattern; the pattern distinctly varies from one individual to the other.



Twins

Monozygotic twins have absolute similarities in terms of protein bands.

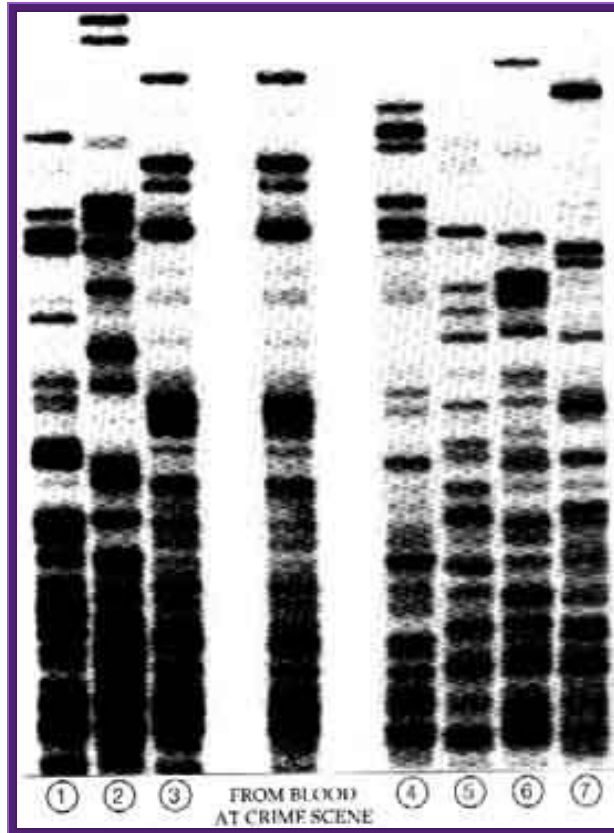
=	=
-	-
-	-
=	=
=	=
=	=
=	=
-	-
-	-
-	-
=	=
-	-
=	=
=	=
=	=



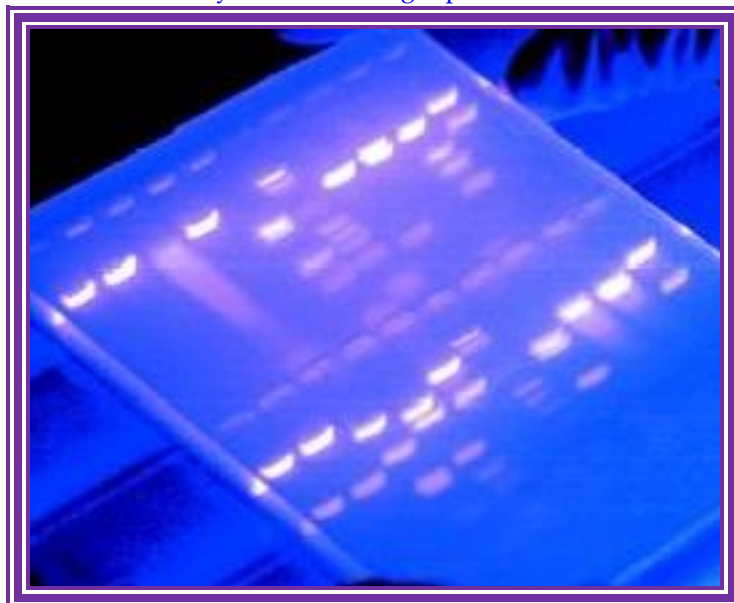
Brother and sister:



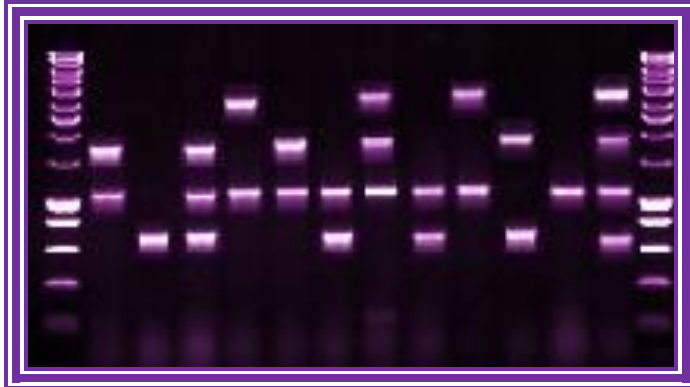
In this case there are many common bands for they are inherited from their parents.



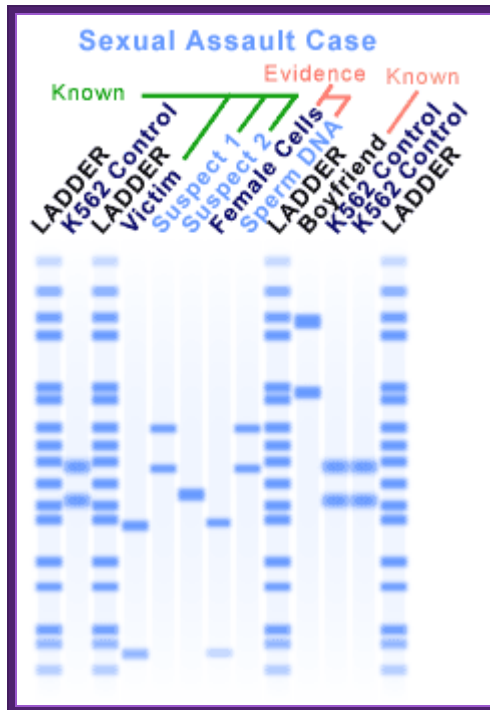
Analysis of DNA finger prints.



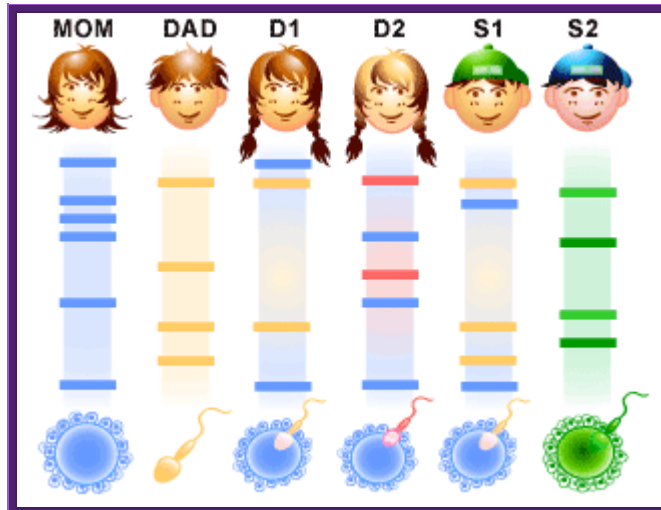
Genetic finger printing-Gel analysis



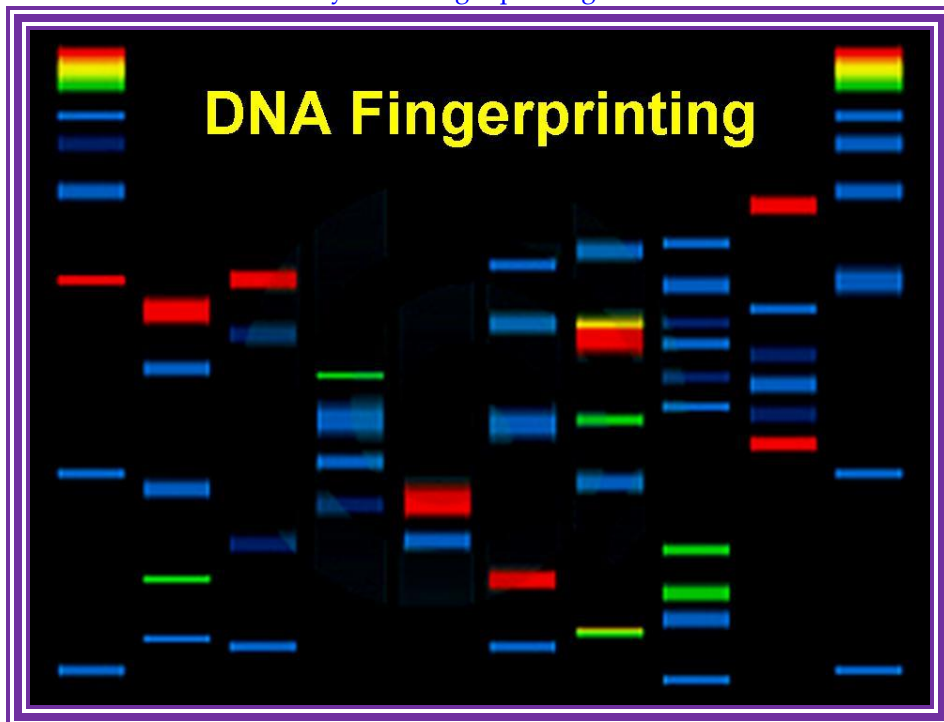
Genotyping



Sexually assaulted cases



Family DNA finger printing



DNA finger prints

Polymorphic DNA refers to chromosomal regions that vary widely from individual to individual. By examining several of these regions within the genomic DNA obtained from an individual, one may determine a "DNA Fingerprint" for that individual. DNA polymorphisms are now widely used for determining

paternity/maternity, kinship, identification of human remains, and the genetic basis of various diseases. The most far-reaching application, however, has been in the field of criminal forensics. DNA from crime victims and offenders can be matched to crime scenes, often affecting the outcome of criminal and civil trials.

Microsatellites, also known as **Simple Sequence Repeats (SSRs)** or **short tandem repeats (STRs)**, are repeating sequences of 2-6 **base pairs** of **DNA**.^[1]

Microsatellites are typically **co-dominant**. They are used as **molecular markers** in **genetics**, for **kinship**, **population** and other studies. They can also be used to study **gene duplication** or **deletion**. Microsatellites are also known to be causative agents in human disease, especially neurodegenerative disorders and cancer.

For example, the following sample report from this commercial DNA paternity testing laboratory Universal Genetics signifies how relatedness between parents and child is identified on those special markers:

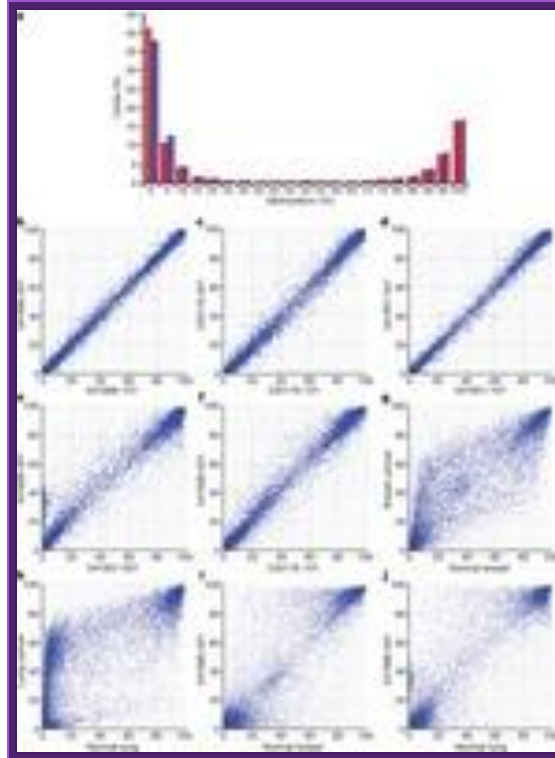
DNA Marker	Mother	Child	Alleged father
D21S11	28, 30	28, 31	29, 31
D7S820	9, 10	10, 11	11, 12
TH01	14, 15	14, 16	15, 16
D13S317	7, 8	7, 9	8, 9
D19S433	14, 16.2	14, 15	15, 17

The partial results indicate that the child and the alleged father's DNA match among these five markers. The complete test results show this correlation on 16 markers between the child and the tested man to draw a conclusion of whether or not the man is the biological father.

Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis: Sergio E. Baranzini et.al.

Monozygotic (MZ) or “identical” twins have been widely studied to dissect the relative contributions of genetics and environment in human diseases. In multiple sclerosis (MS), an autoimmune demyelinating disease and common cause of neurodegeneration and disability in young adults, disease discordance in MZ twins has been interpreted to indicate environmental importance in its pathogenesis¹⁻⁸. However, genetic and epigenetic differences between MZ twins have been described, challenging the accepted experimental paradigm in disambiguating effects of nature and nurture. Here, we report the genome sequences of one MS-discordant MZ twin pair and messenger RNA (mRNA) transcriptome and epigenome sequences of CD4⁺ lymphocytes from three MS-discordant, MZ twin pairs. No reproducible differences were detected between co-twins among ~3.6 million single nucleotide polymorphisms (SNPs) or ~0.2 million insertion-deletion polymorphisms (indels). Nor were any reproducible differences observed between siblings of the three twin pairs in HLA haplotypes, confirmed MS-susceptibility SNPs, copy number variations, mRNA and genomic SNP and indel genotypes, or expression of ~19,000 genes in CD4⁺ T cells.

Figure.: Comparisons of methylation of genomic CpG sites in CD4⁺ lymphocytes and breast and lung tissue samples.



a, Frequency distribution of CpG site methylation in 041896-001 (blue) and -101 (red) using ELAND-extended. **b–j**, Pair-wise comparisons of CpG site methylation using ELAND-extended in CD4⁺ lymphocytes from monozygotic twin siblings 041896-001 and -101

Only two to 176 differences in methylation of ~2 million CpG dinucleotides were detected between siblings of the three twin pairs, in contrast to ~800 methylation differences between T cells of unrelated individuals and several thousand differences between tissues or normal and cancerous tissues. In the first systematic effort to estimate sequence variation among MZ co-twins, we did not find evidence for genetic, epigenetic or transcriptome differences that explained disease discordance. These are the first female, twin and autoimmune disease individual genome sequences reported.

DNA methylation profiles in monozygotic and dizygotic twins:

Zachary A Kaminsky et.al,

Twin studies have provided the basis for genetic and epidemiological studies in human complex traits^{1,2}. As epigenetic factors can contribute to phenotypic outcomes, we conducted a DNA methylation analysis in white blood cells (WBC), buccal epithelial cells and gut

biopsies of 114 monozygotic (MZ) twins as well as WBC and buccal epithelial cells of 80 dizygotic (DZ) twins using 12K CpG island microarrays^{3,4}. Here we provide the first annotation of epigenetic metastability of 66,000 unique genomic regions in MZ twins. An intraclass correlation (ICC)-based comparison of matched MZ and DZ twins showed significantly higher epigenetic difference in buccal cells of DZ co-twins ($P = 1.2 \times 10^{-294}$). Although such higher epigenetic discordance in DZ twins can result from DNA sequence differences, our in silico SNP analyses and animal studies favor the hypothesis that it is due to epigenomic differences in the zygotes, suggesting that molecular mechanisms of heritability may not be limited to DNA sequence differences. (Buccal cavity differences are local but not inherited).

DNA Finger Prints: Technique of finger printing was perfected by British **genetist Dr. Alec Jeffreys in 1984.**



Fig. above, Identical Twins: DNA finger printing of identical and monozygotic twins did not show any variations in their DNA sequences.