



# Molecular Genetics of Human Disease

*Hesham Abdeen*  
*Belfast City Hospital*



## Molecular Genetics ↔ Industrial Biotechnology

- *The advent of high throughput genome and gene/protein expression analyses has transformed our understanding of biological systems and enabled us to control and exploit these systems for our benefit*
- *The wealth of knowledge accumulated over the last few years on the structure of genomes of many organisms and how they operate has opened the horizon for unlimited applications in almost every aspect of life*
- *There is a wide range of applications in three vital sectors, including:*
  - *Health care/pharmaceutical sector*
  - *Agro-food sector*
  - *Industrial manufacturing, energy and environment*



## Molecular Genetics ↔ Health Care/Pharmaceuticals

- *The functional fitness of the genome contributes significantly to the function/malfunction and longevity of the various organs of the body.*
- *Almost 60% of the Human population is thought to have a genetically-influenced disease representing a major impact on the economy and health services all over the world*
- *Identification of genes implicated in genetic and genetically-influenced diseases facilitates early prediction of risk, diagnosis and provides hope for the development of drug- or gene-based therapies*

## Molecular Genetics ↔ Health Care/Pharmaceuticals

- ***Pharmaco-genetics***: development of novel therapies and to improve efficacy and reduce side effects of therapies already in use.
- ***Immuno-genetics***: analysis of genetic susceptibility and resistance to parasitic infections may lead to the development of effective interventions and/or vaccines
- ***Genetic testing and histo-compatibility***
- ***Forensic science and historical studies***

---

In genetic testing, there are well established genetic tests that could diagnose accurately and cheaply hundreds of diseases

Immunogenetics is rapidly expanding field of science. Almost 10% of the Human genome is involved in self-recognition and immunity against parasitic infections such as Hepatitis C virus. Analysis of genes such as the HLA and KIR gene clusters may provide a better understanding of .the mechanisms of susceptibility and resistance to parasitic infections

Furthermore the matching process of donors and recipients in organ transplantation relies completely on molecular analysis of the HLA class I and II genes

Substitutive therapies such as insulin and Interferon have been in use for decades. The latest knowledge of the gene clusters involved in production and modulation of the activity of these therapies could be now invested in the development of more efficient safer treatments

Molecular analysis of Human remains can now be used to resolve the mystery of crimes and to ID ancient remains such as mummies. We probably all know about the latest discoveries of Dr. Hawas in this field. And now we have the technology to be able to ID 10 of thousands of mummies in a matter of a couple of years, rewriting the Egyptian history



# Molecular Genetics of Human Disease

- *Analysis of the genetic background of any Human disease remains a challenging task*
- *Molecular genetics represents a set of powerful tools which utilize the latest updates of our understanding of the Human genome structure to zoom in on genomic regions harboring disease-causing mutations*

---

.The genetic make up of a person makes alot of difference between fitness and illness

s26

It is thought that each one of us carries on average 4 genetic defects ranging from mild vision and hearing problems to more sever diseases

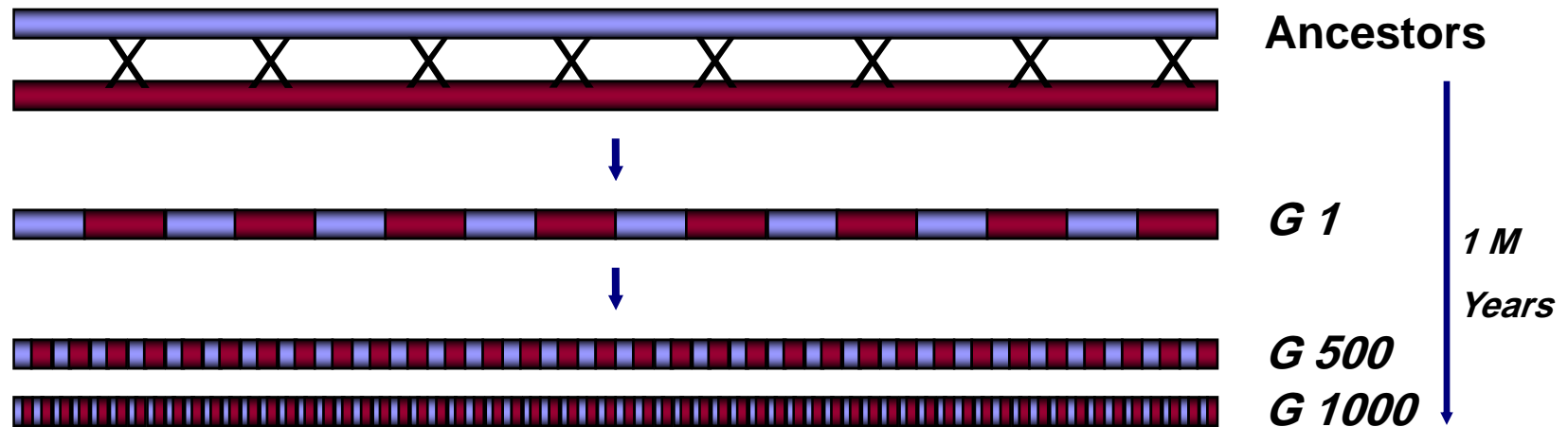
The Human genome is 3 BN nucleotides long. Unfortunately, we can not sequence the entire genome in affected individuals at an affordable  
-----  
.budget

.Molecular genetics uses the latest updates of our understanding of the genome struture to zoom in on genomic regions harboring mutations

ssss; 01/02/2010

# Evolution of the Human Genome

## ➤ Recombination



---

.First lets cover some of the basic discoveries which allowed the study of such a complex system

s1

Since the publication of the draft sequence of the Human genome and subsequent in depth analysis, we have realized that the Human genome is a dynamic ever evolving system. This evolution process is mediated and controlled biological, environmental and social factors shaping .Human communities all over the world

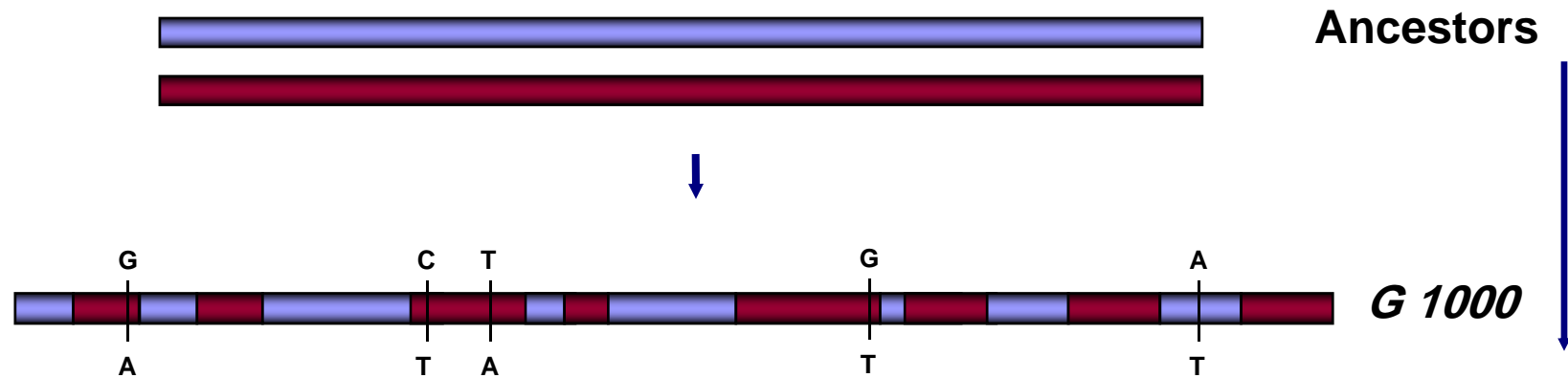
Recombination is one of the primary process driving the Human genome's evolution. Since our ancestors walked the earth, their genomes have been continuously blending together through crossing over\_\_\_\_\_ of their chromosomes. There are on average 50 recombinations in every new .borne baby

.After one million years of evolution our genome has become a mosaic of bits and peactices of our ancestors' genomes

ssss; 01/02/2010

# Evolution of the Human Genome

## ➤ Mutation



## ➤ Environmental and Social factors

- Environmental: weather, epidemics, natural barriers, natural disasters, .....
- Social: customs, religion, ethnic origin, historical events, .....

Mutation is another key biological process involved in the evolution of the Human genome. Genetic materials rarely suffer mutations in the form  
 .(represented in this schematic by the distorted banding pattern)of single base changes , insertions, deletions and duplications

s3

Many of these mutations have no or mild effect on the performance of the genome, while others are pathogenic or even fatal. Although  
 mutations occur at an extremely low rate, their accumulation over time through thousands of generations resulted in a wide variety of  
 .functional variants and genetic disorders

The evolution of the human genome is also regulated by a number of environmental and social factors which enrich or reduce the heterogeneity  
 of our genetic pools

:These factors include

:Environmental factors

Weather the driving force of natural selection

Epidemics has wiped out individuals with genomes unfit to confer resistance against parasitic infections such as influenza, measles, yellow  
 .fever, plague and malaria

Natural barriers prevent free migration and admix among populations

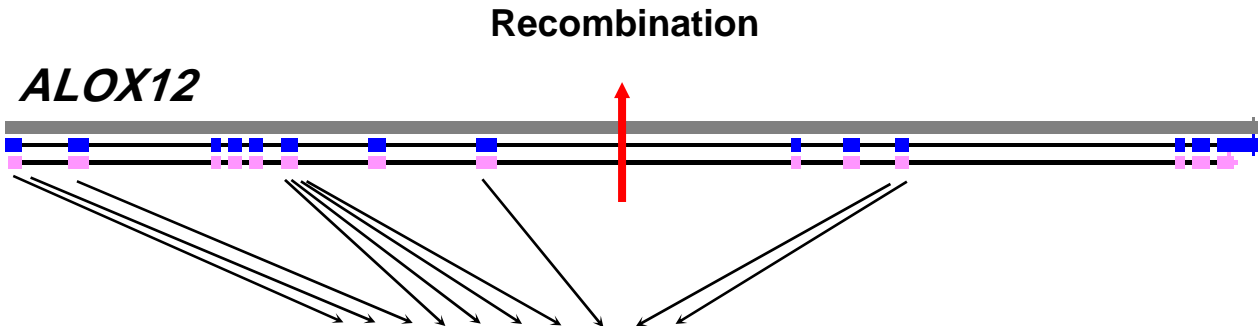
Social factors such as customs, religion and ethnic origin, have also placed restrictions over mixing of Human communities

Finally historical events, such as large-scale wars wiped out entire communities or triggered mass migrations

we all probably studied in history the campaign of the mongols which probably altered the genetic make up of the Asian and Middle Eastern  
 .populations for ever

ssss; 01/02/2010

# Molecular Genetics Exercise



		Block A					Block B		Haplotype			
		1	2	3	4	5	6	7	8	9	10	
1	CACGAAAG	A	C									1
2	CACGAAAG	G	C									2
3	CACGAAAG	A	T									3
4	TGTAGAGA	A	T									4
5	TGTAGAGA	A	C									5
6	CGTAGTGA	G	C									6

- Haplotypes
- Haplotype Blocks
- Linkage
- Haplotype Tagging (HT)

As I mentioned earlier, the Human genome has gained through recombination a block-like structure representing a mosaic of short segments of our ancestors' genomes

.This is an example of one of the genes I have been working on named ALOX12

.The gene spans 14 exons which contain 10 SNPs

Analysis of 100 individual chromosomes revealed 6 distinct haplotypes \_\_\_\_\_of the gene

.Close examination of these haplotypes revealed the existence of two major haplotype blocks \_\_\_\_\_A and B

Any of the sequences of block A can go with any of the the sequences of block B. This indicates a historical recombination event which splitted .the gene into two blocks

For example the CACGAAAG sequence of block A can go with any of the AC, GC or AT sequences of Block B

Alleles in each block appear to be in Linkage \_\_\_\_\_with each others

.as you can see here the A allele of SNP2 is linked with the C allele of SNP3 while the G allele is linked with the T allele

Due to low sequence heterogeneity\_\_\_\_\_with in each block, the knowledge of the genotypes of a few SNPs should be enough to deduce the .rest of the haplotype sequence

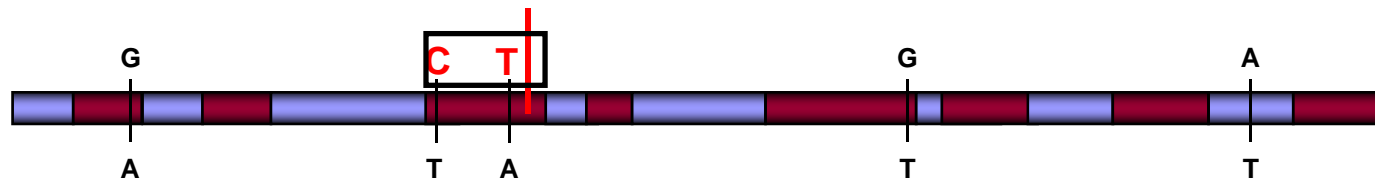
Rather than the expensive time consuming option of having to sequence all 14 exons of the gene. genotyping SNPs\_\_\_\_\_4, 6, 9 and 10 .should be enough to distinguish between the 6 different haplotypes of the gene

For example, if there is a G in position 4 with an AC on positions 9 and 10 you would know you that the individual contains haplotype 1, while .if it goes with a GC then the individual contains haplotype 2, and so on

This economic genotyping strategy is called \_\_\_\_\_Haplotype-tagging and it represents the corner stone of all recent molecular genetics .studies

# Whole-Genome Screening

Unknown Pathogenic Mutation



SNP	$p$ -value
G/A	0.178
<b>C/T</b>	<b>0.0004</b>
<b>T/A</b>	<b>0.000009</b>
G/T	0.778
A/T	0.981

← Polymorphisms linked to Pathogenic Mutation

---

The discovery of haplotype blocks containing alleles tightly linked with each others has placed the foundation of modern molecular genetics

s9

Suppose there is a disease causing mutation somewhere on the genome. Then this mutation is expected to be linked with a haplotype block containing certain alleles of adjacent polymorphisms

These alleles are expected to be overrepresented in affected individuals than the rest of the normal population

So in theory if a sufficient number of haplotype tagging polymorphisms is genotyped to cover the entire genome (whole-genome screening), statistical analysis should identify the region containing alleles overrepresented in patients

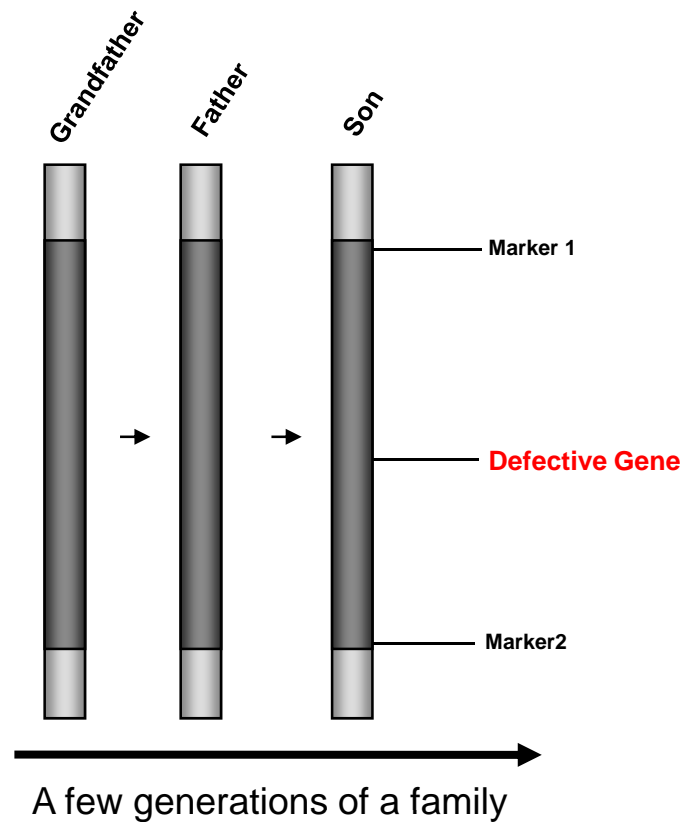
.This region is highly likely to harbor the disease causing mutation

Full sequence analysis of the genes in such region should lead to the identification of the pathogenic mutation

ssss; 02/02/2010

# Analysis of The Genetic Backgrounds of Human Diseases

## I. Linkage Analysis in Affected Families



---

.There are two major approaches to carry out whole-genome genetic studies, linkage and association analysis

s10

.Both approaches search for genetic markers which segregate in affected more than normal individuals

(single-gene disorders)Linkage analysis is mainly carried out to study Human diseases of simple background

.Single gene disorders have a familial pattern of inheritance and usually looked at in families with several affected relatives

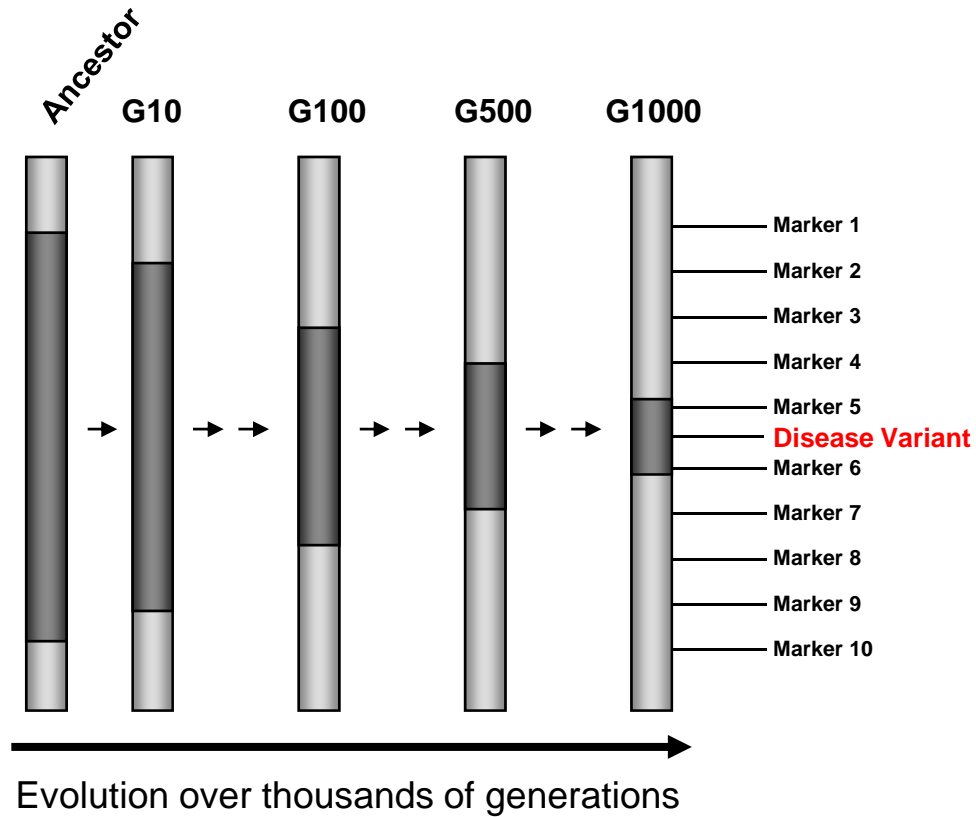
In a family the disease gene is expected to be co-inherited along with other normal genes and genetic markers on a large block, which remains relatively unchanged in a very few number of generations, from grandfather to father to son

.A widely-spaced panel of a few hundred makers is usually sufficient to localize the region harboring the defective gene

ssss; 01/02/2010

# Analysis of The Genetic Backgrounds of Human Diseases

## II. Association Analysis in Populations



---

On the other hand, association analysis is mainly carried out to study multi-gene disorders with complex genetic background

s11

These disorders such as hypertension and diabetes have no obvious familial pattern of inheritance and are usually analysed in a population rather than families

However, the genomic blocks containing the disease causing mutations are usually reduced in size through the evolution of the population over thousands of generations

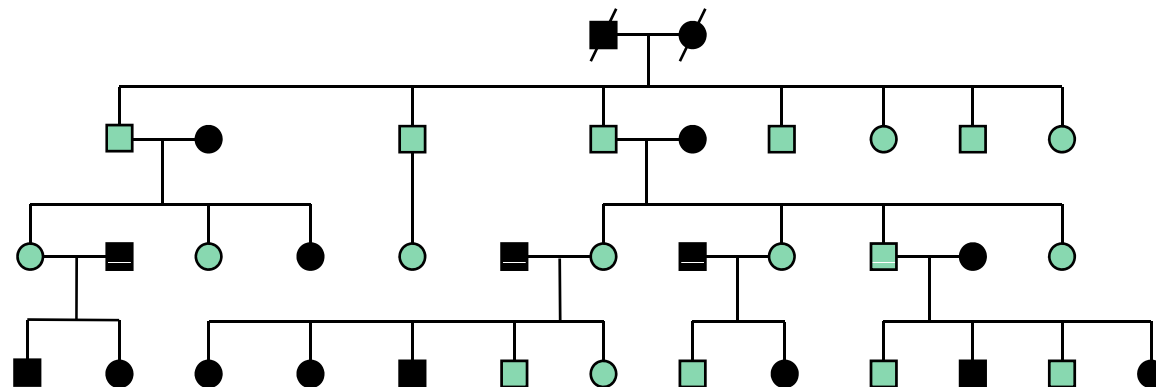
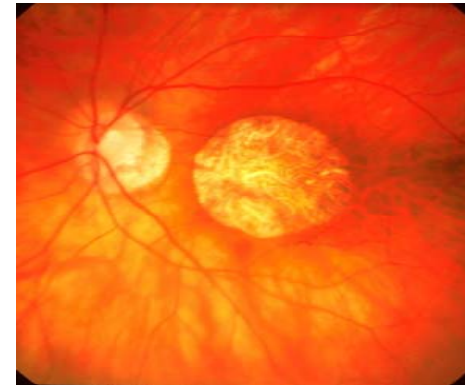
.A tightly-spaced panel of hundreds of thousands of markers is required to scan the genome for association with a disease

.Furthermore, to achieve sufficient statistical power, association studies need to include thousands of patients and controls

ssss; 02/02/2010

# Central Areolar Choroidal Dystrophy (CACD)

- *CACD is a macular degenerative disease causing loss of vision in middle age*
- *Mode of inheritance is autosomal dominant*



Over the next section of this talk I'm going to describe some of our work on a single-gene and multi-gene disorders. First the single gene disorder CACD  
ssss; 01/02/2010



## CACD Mutation Maps to Chromosome 17p13.2

- *The CACD mutation maps to a 2.0 Mb block on chromosome 17p13.2*
- *The same region also links to dominant cone dystrophy (CORD5) and contains two mutations involved in recessive Leber congenital amaurosis*
- *It is also a hot locus for more than 15 different types of cancers*

## Haplotype Sharing Among Families and Sporadic Cases

Marker	Families					Sporadic Cases			
	1	2	3	4	5	1	2	3	4
D17S513	2	1	1	3	4	1	1	3	3
<i>KIAA0523</i>	1	3	3	8	7	3	3	7	7
D17S938	2	2	1	1	1	2	2	3	3
<i>AIPL1</i>	4	4	5	5	5	4	4	5	5
<i>FLJ10156</i>	1	1	1	1	1	1	1	1	1
<i>NIR1</i>	2	1	4	4	3	1	1	4	4
D17S1881	4	4	2	4	3	4	2	1	5
D17S906	3	2	4	2	1	1	3	6	5
D17S578	1	1	3	1	1	1	2	1	1
<i>ALOX12</i>	1	1	1	2	3	1	1	1	1
<i>ASGR1</i>	2	2	1	3	4	1	1	3	3
<i>DLG4</i>	3	3	3	1	4	3	3	2	2
<i>YBX2</i>	2	4	2	2	3	2	5	1	1
D17S960	2	2	1	2	2	2	2	2	2
D17S729	1	1	3	1	1	1	1	1	1
<i>FGF11</i>	1	3	1	3	1	3	2	1	1
<i>TNFSF12</i>	1	4	1	2	2	3	1	1	2

} 200 kb

---

.Microsatellite and HT SNP markers were genotyped in 5 affected families and 4 sporadic cases

s16

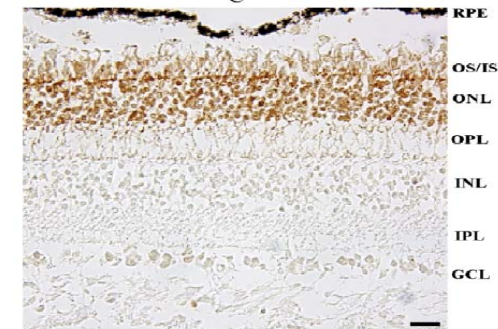
Comparison of haplotypes revealed substantial similarities in the distal part of the region including a haplotype block of 200 kbs which is shared among all families and cases

There are two distinct haplotypes (represented in green and orange) indicating the existence of two different mutations leading to CACD

ssss; 02/02/2010

# Aryl Hydrocarbon Receptor-interacting Protein-Like 1 (AIPL1)

- *AIPL1 is one of the genes involved in recessive leber congenital amaurosis*
- *Uniquely expressed in the retina*
- *Plays an important role in the regulation of cell cycle during photoreceptor maturation and subsequent survival*
- *Involved in key functions in the retina and thus represents an excellent candidate for the CACD mutation*



.Fortunately, the 200 kb CACD block contains one gene named AIPL1

s17

This gene is the subject of full sequence analysis using Next-generation sequencing at the moment and we are hoping to identify the CACD .mutation very soon

ssss; 01/02/2010



# Complex Genetic Autoimmune Diseases

- *Autoimmune diseases affect ~5% of the human population*
- *They are believed to arise from immune-mediated attack against self-antigens in specific organs and tissues*
- *Substantial evidence from epidemiological studies indicates the involvement of both environmental and multi-genetic components*
- *It appears that a common patho-physiological mechanism is involved in predisposition to several autoimmune diseases*

.Well it seems linkage analysis works very well for a single gene disorder

s13

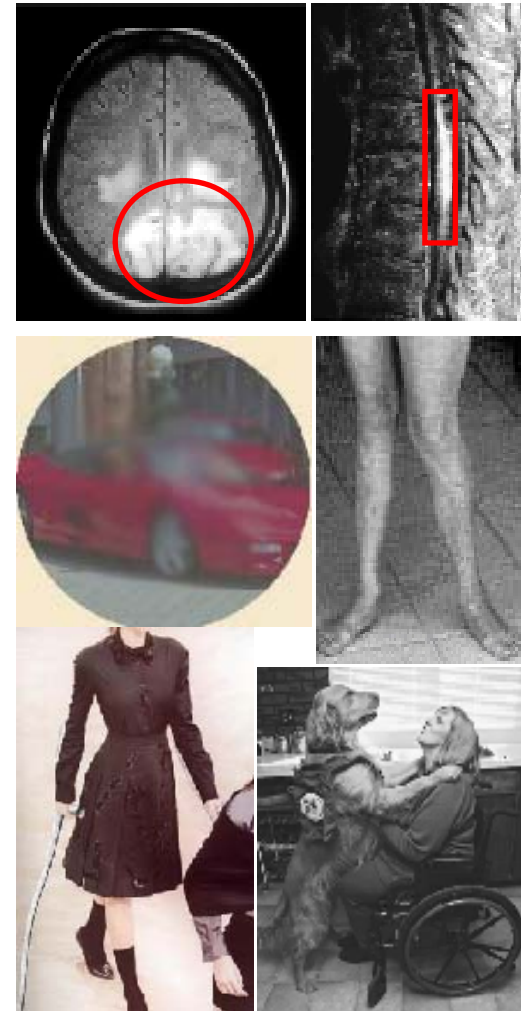
Lets see if an association study is capable of resolving the mistry of a multi-geneic complex disease

.In many cases affected individuals tend to develop more than one autoimmune disease

ssss; 02/02/2010

# Multiple Sclerosis

- *Multiple sclerosis (MS) is a common autoimmune inflammatory disease of central nervous system*
- *Common symptoms include sensory and visual disturbances, loss of balance and coordination, bladder and bowel incontinence, pain, weakness, fatigue and paralysis*
- *The disease hits patients in their early adulthood and its two to three times more frequent in women than men*
- *Despite the severity of the symptoms, the life span of affected individuals is only slightly shortened creating a significant impact on the quality of their life*



.characterized by demyelination within the central nervous system-----

s18

Common symptoms include sensory and visual disturbances, loss of balance and coordination, bladder and bowel incontinence, pain,  
.weakness, fatigue and paralysis

ssss; 02/02/2010



## MS Genetics

- *In MS, whole-genome association studies have been carried out in different populations from 18 countries in North America, Europe and Australia*
- *Most studies indicated the presence of a strong susceptibility locus in the HLA region on chromosome 6p21*
- *However, there was little agreement about other suggested regions*
- *An updated meta-analysis of all association screens was carried out to verify findings, and draw a better image of MS genetic background*

In other words, findings in one population are not necessary true in another

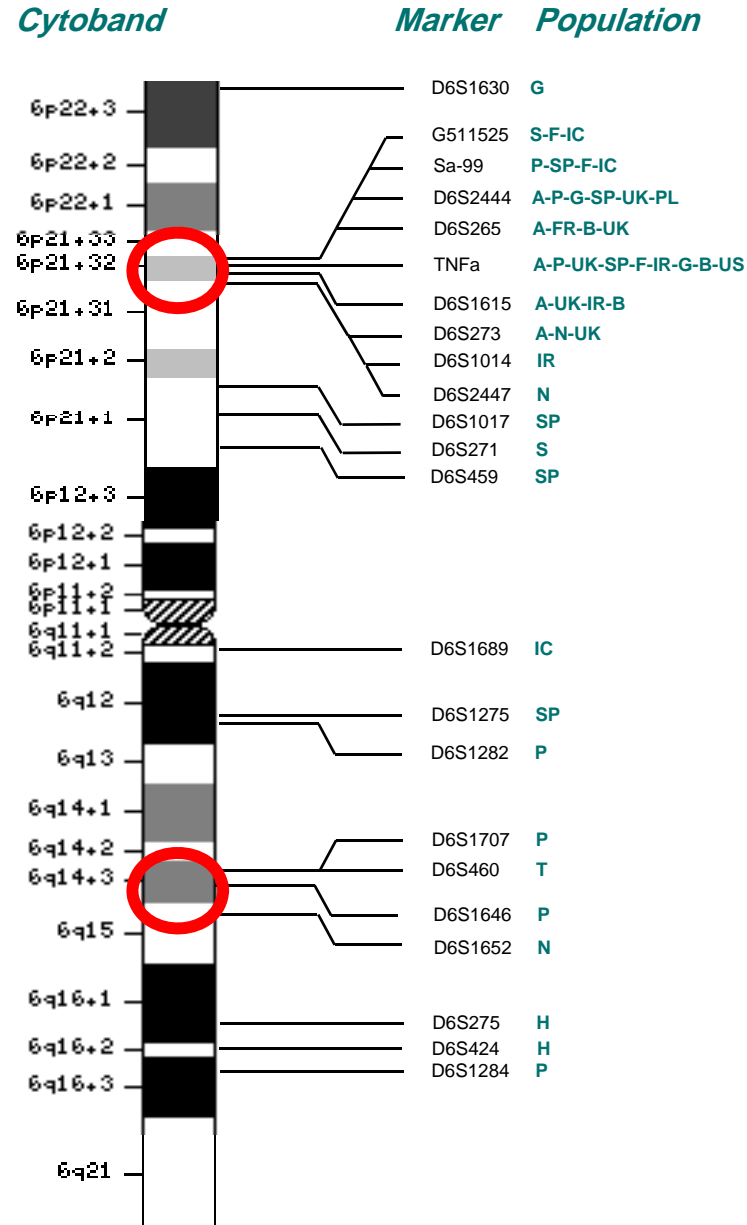
ssss; 02/02/2010

# Meta-analysis

- A**      **Australia**
- B**      **Belgium**
- F**      **Finland**
- FR**     **France**
- G**      **Germany**
- IC**     **Iceland**
- IR**     **Ireland**
- I**      **Italy**
- N**      **Nordic**
- P**      **Portugal**
- PL**     **Poland**
- S**      **Sardinia**
- SP**     **Spain**
- T**      **Turkey**
- UK**     **United Kingdom**
- US**     **United States**

**6p21**

**6q14**



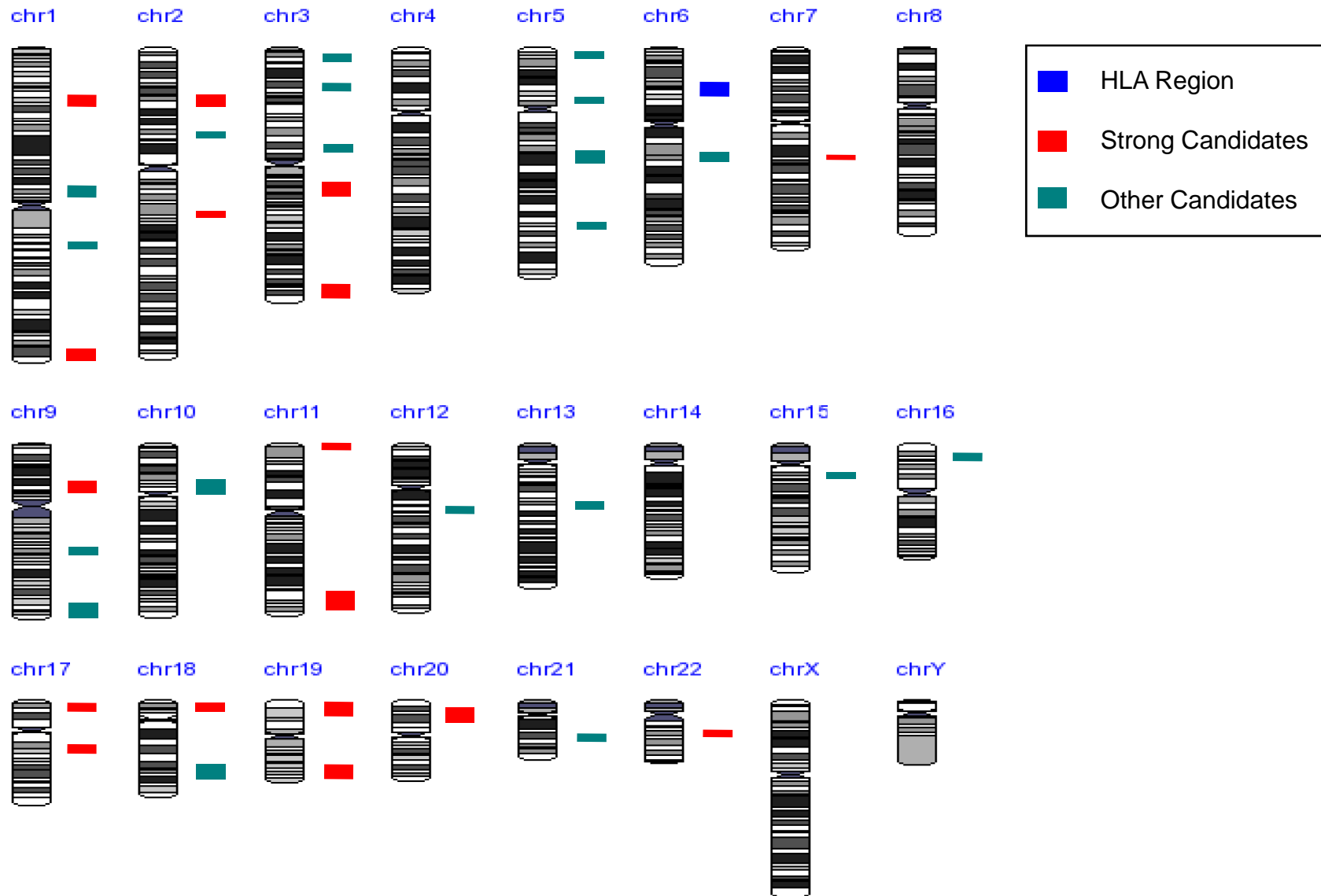
This meta-analysis was carried out by collating findings from all whole genome screens in a single genomic map showing the locations of every genetic marker reported in the list of populations on the left-hand side

s20

The map was screened using a sliding window searching for blocks consistently reported through out several populations

ssss; 01/02/2010

# MS Genetic Susceptibility Map



---

In addition to the HLA region, the map predicted at least 37 blocks harboring possible susceptibility genes, including a number of novel candidates such as chromosomes 1p13 and 9p22

s22

Of these, 17 regions (shown in red) have been reproducibly reported in many populations, thus representing Strong candidate regions for MS susceptibility

Interestingly, the map demonstrated major similarities between MS and other autoimmune diseases. 15 of the possible MS susceptibility loci have been previously identified in other autoimmune diseases such as type 1 diabetes mellitus (IDDM), rheumatoid arthritis (RA) and systemic (SLE)lupus erythematosus

ssss; 02/02/2010



## Genomic Screening of The N. Irish Population

- *Detection of most previously reported susceptibility regions to MS and other autoimmune diseases indicates the ability of this MS map to predict genomic regions harboring susceptibility genes*
- *The map excluded almost 95% of the Human genome allowing detection of MS susceptibility regions with massive reduction of cost and time.*
- *This ability of the map to predict MS susceptibility was further tested in a genomic screen of the N. Irish population*

# Regions Showing Association in The N. Irish

- *16 Genomic regions showed association,*

---

## *Cytoband P-values Candidate genes*

---

20p12	0.0006	<i>BMP2, PAK7, pFUSIP1</i>
1p13	0.0008	<i>CD53, CSF1, NGFB, PTPN22</i>
2q14	0.0008	<i>CNTNAP5</i>
7q21	0.0011	<i>ADAM22, GRM3</i>
2p13	0.0013	<i>CNB1, TGFA</i>
17p13	0.0014	<i>NTN1, TNFRSF12, TNFRSF13</i>
15q13	0.0019	<i>CHRNA7</i>
11p15	0.0020	<i>TH, CTSD</i>
5q14	0.0035	
13q14	0.0040	<i>THSD1</i>
3p23	0.0089	<i>CKLFSF6, 7 and 8, TGFBR2</i>
5p13	0.0102	<i>C6, C7, C9, GDNF, IL-7R</i>
19p13	0.0102	<i>ADAMTS10, ICAM1, 3, 4 and 5, OR7E</i>
18q21	0.0112	<i>TNFRSF11A</i>
11q23	0.0363	<i>DRD2, DSCAML1, IL-18, IL10-RA, NCAM1</i>
3q13	0.0450	<i>DRD3, GAP43, LSAMP, NP25</i>

---



## Candidate Genes

- *Most of these regions contain genes with key roles in neuronal development, survival and regeneration that could be easily involved in susceptibility to MS*
- *There are also a number of genes encoding immunoregulatory and signaling proteins, which may contribute to the immune responses observed through the disease course*
- *To verify candidate genes, many of the confirmed regions are the currently subjected to high-resolution association analysis*



## Conclusion

- *A universal genomic map of possible susceptibility regions to MS was constructed*
- *The map revealed several strong possible regions and emphasized the role of a number of regions among several autoimmune diseases such as MS, IDDM and RA*
- *Using this map, we were able to determine the genetic background of MS in the N. Irish at a greatly reduced cost and time*
- *This confirms that the map in deed provides a comprehensive view of genetic susceptibility to MS, a first step towards the identification of the genes implicated in the disease*



***Colin Graham & Shirley Heggarty***

**Molecular Genetics, Belfast City Hospital**

***Stanley Hawkins, Aidan Droogan & Gavin McDonnell***

**Department of Neurology, Royal Victoria Hospital, Belfast**

***Michael Hutchinson***

**Department of Neurology, St Vincent's Hospital, Dublin**

***Stephen Sawcer & Alistair Compston***

**Department of Neurology, Addenbrooke's Hospital, Cambridge**

**This work was supported with generous funding from:**

**The Guide Dogs for The Blind Association**

**Genetic Analysis of MS in EuropeanS Project (GAMES)**

**MS Ireland**