Human Genetic Polimorphism

Human genome

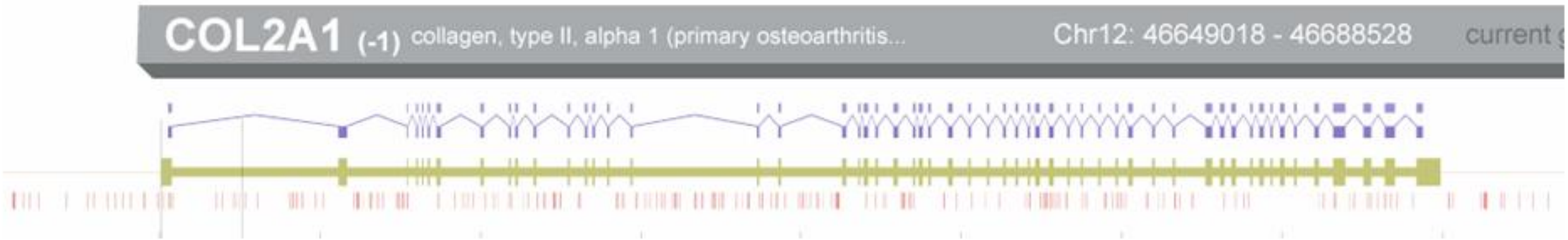
- 3.1 billion (10^9) base pairs
- 22,500 genes
- encoding 100,000 proteins

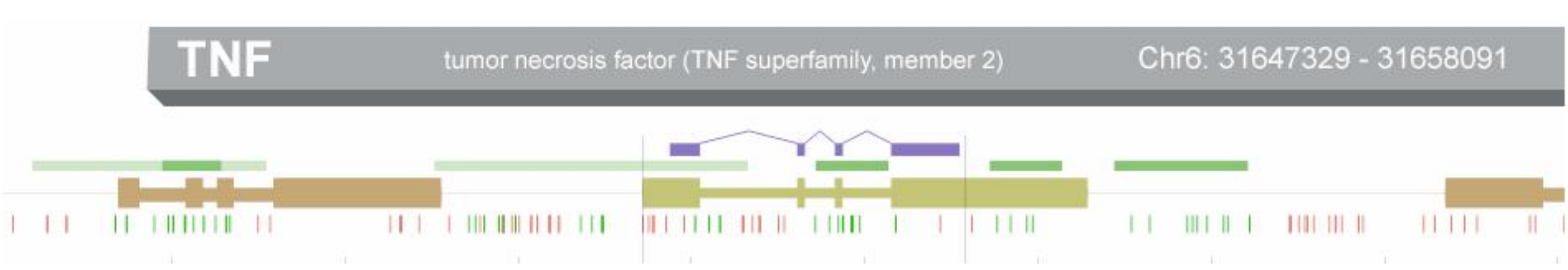
SNP – **s**ingle **n**ucleotide **p**olymorphisms

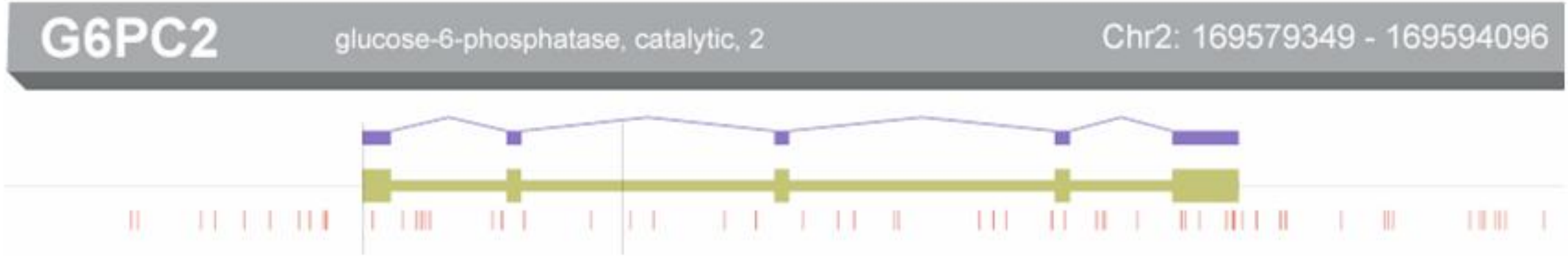
<u>3 million or 0.1%</u> bases in our DNA are polymorphic as single nucleotide polymorphisms (SNP).

Directly responsible for majority of variation among humans

SNP (**s**ingle **n**ucleotide **p**olymorphisms) examples

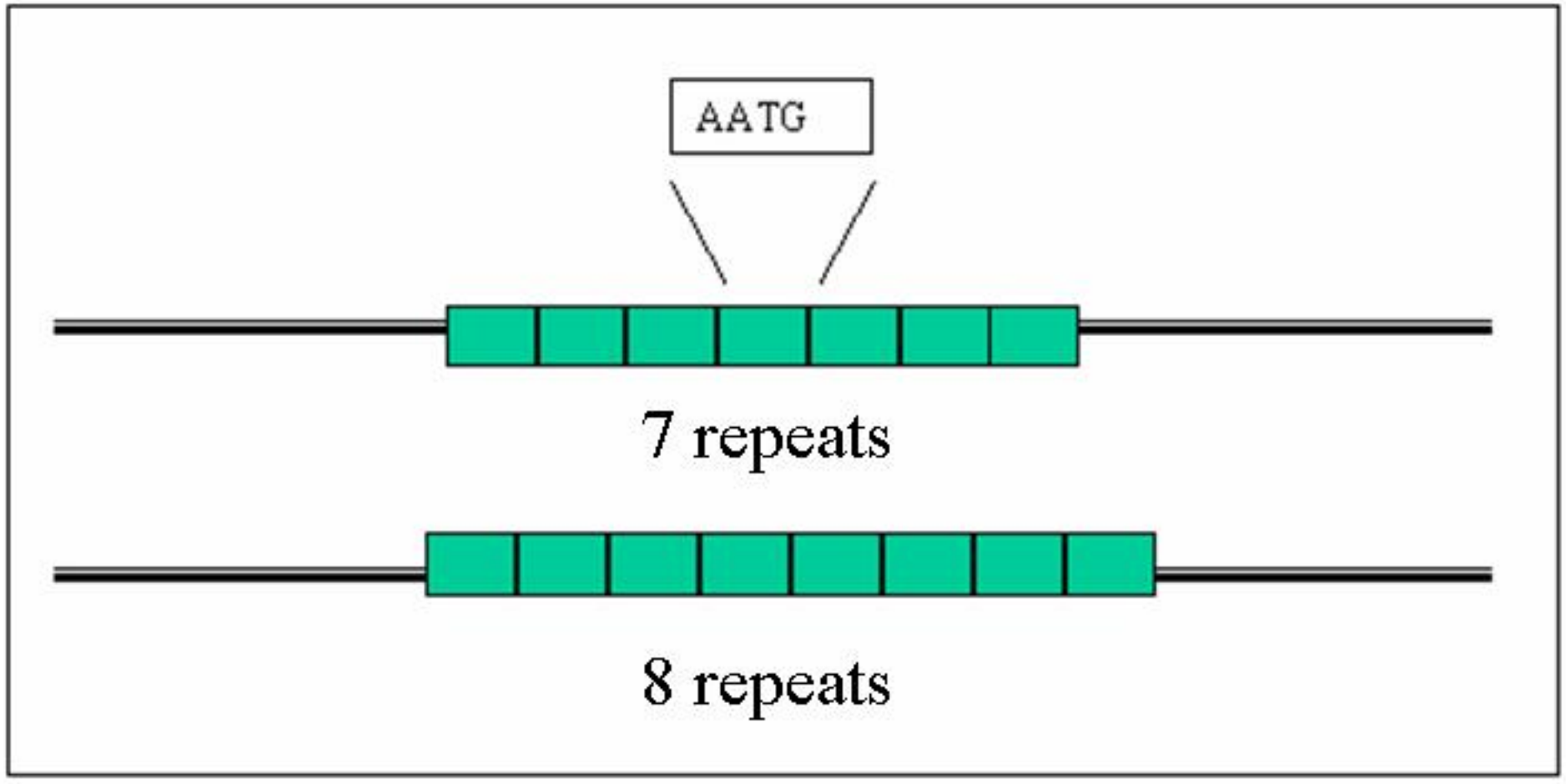






STR – **S**hort **T**andem **R**epeats

STR (short tandem repeats) structure



STRs vs. SNPs

Marker	Advantages	Disadvantages
SNPs	• Low mutation rate • High abundance • Easy to type • New analytical approaches are being developed at present • Cross-study comparisons are easy; data repositories already exist	• Substantial rate heterogeneity among sites • Expensive to isolate • Ascertainment bias • Low information content of a single SNP
Microsatellites	• Highly informative (large number of alleles, high heterozygosity) • Low ascertainment bias • Easy to isolate	• High mutation rate • Complex mutation behaviour • Not abundant enough • Difficult to automate • Cross-study comparisons require special preparation

STRs are important genetic markers (first maps of human genome)

The physiological function is not certain

but some STRs are directly involved in pathogenesis

Poly Glutamine (Q) triplet repeat diseases

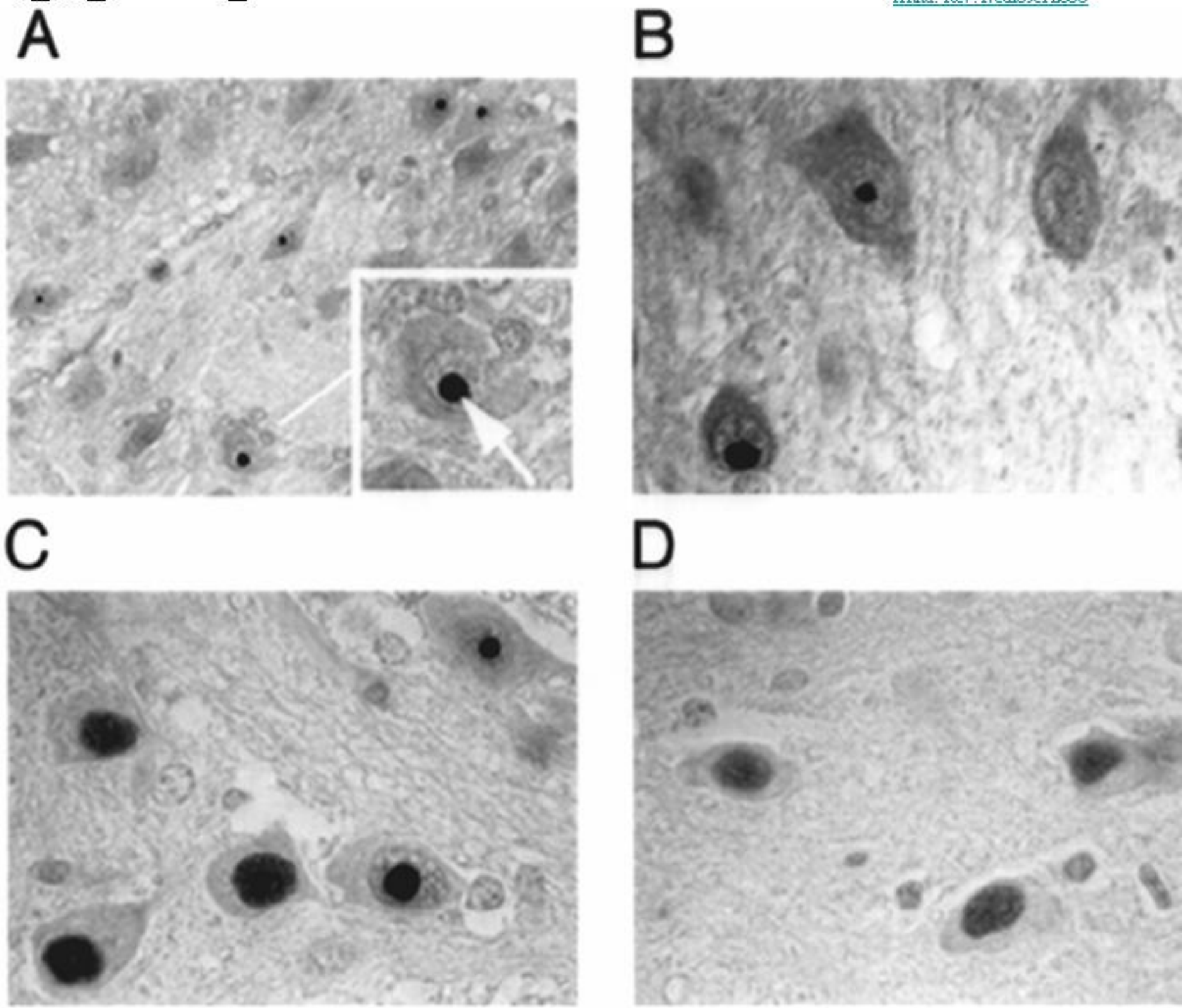
Disease	Repeat sequence
HD (Huntington Disease)	CAG
DRPLA (dentatorubral pallydoluysian atrophy)	CAG
SCA (spinocerebellar atrophy) 1, 2, 3, 6, 7, 17	CAG
Kennedy's disease	CAG

Brain 2004

Polyglutamine stretches predispose to aggregate formation

Annu. Rev. Neurosci 2000

- (*A*, *B*) The subcellular distribution of ataxin-1 in nucleus pontis centralis neurons from a spinocerebellar ataxia 1 patient. Nuclear inclusions (NI) magnified (*bottom right*) and containing ubiquitin (*B*).
- (*C*) Redistribution of the 19S proteasome to aggregates in patient tissue;
- (*D*) control



Conditions and genes in which alanine tract expansion has occurred

Condition	Gene	Gene type	Expansion size	Protein dysfunction
Synpolydactyly type II	HOXD13	Transcription factor	15A -> 22 – 29A (an addition of 7 – 14)	Dominant negative
Cleidocranial dysplasia	RUNX2 (CBFA1)	Transcription factor	17A -> 27A (an addition of 10)	Loss-of-function
Oculopharyngeal muscular dystrophy	PABPN1	Polyadenylate-binding protein	10A -> 11 – 17A (an addition of 1 – 7)	Toxic protein aggregates
Holoprosencephaly (HPE5) Hand-foot-genital syndrome	ZIC2 HOXA13	Transcription factor Transcription factor	15A -> 25A (an addition of 10) 18A -> 24A or 26A (an addition of 6 – 8)	Loss-of-function Unclear, might be dominant negative
Blepharophimosis, ptosis and epicanthus inversus	FOXL2	Transcription factor	14A -> 22 – 24A (an addition of 8 – 10)	Partial loss-of-function
Mental retardation; X-linked, with isolated growth hormone deficiency	SOX3	Transcription factor	15A -> 26A (an addition of 11)	Unknown
Infantile spasm syndrome, X-linked; Partington syndrome; lissencephaly with ambiguous genitalia, X-linked; mental retardation X-linked 36 and 54	ARX	Transcription factor	A-tract #1 (amino acids 100 – 115) 16A -> 18 or 23A (an addition of 2 or 7) A-tract #2 (amino acids 144 – 155) 12A -> 20A (an addition of 8)	Partial loss-of-function
Congenital central hypoventilation syndrome/Ondine curse	PMX2B (PHOX2B)	Transcription factor	20A -> 25 – 29A (an addition of 5 – 9)	Loss-of-function

TRENDS in Genetics 2004

Diseases caused by expansion of <u>non-coding</u> trinucleotide repeats

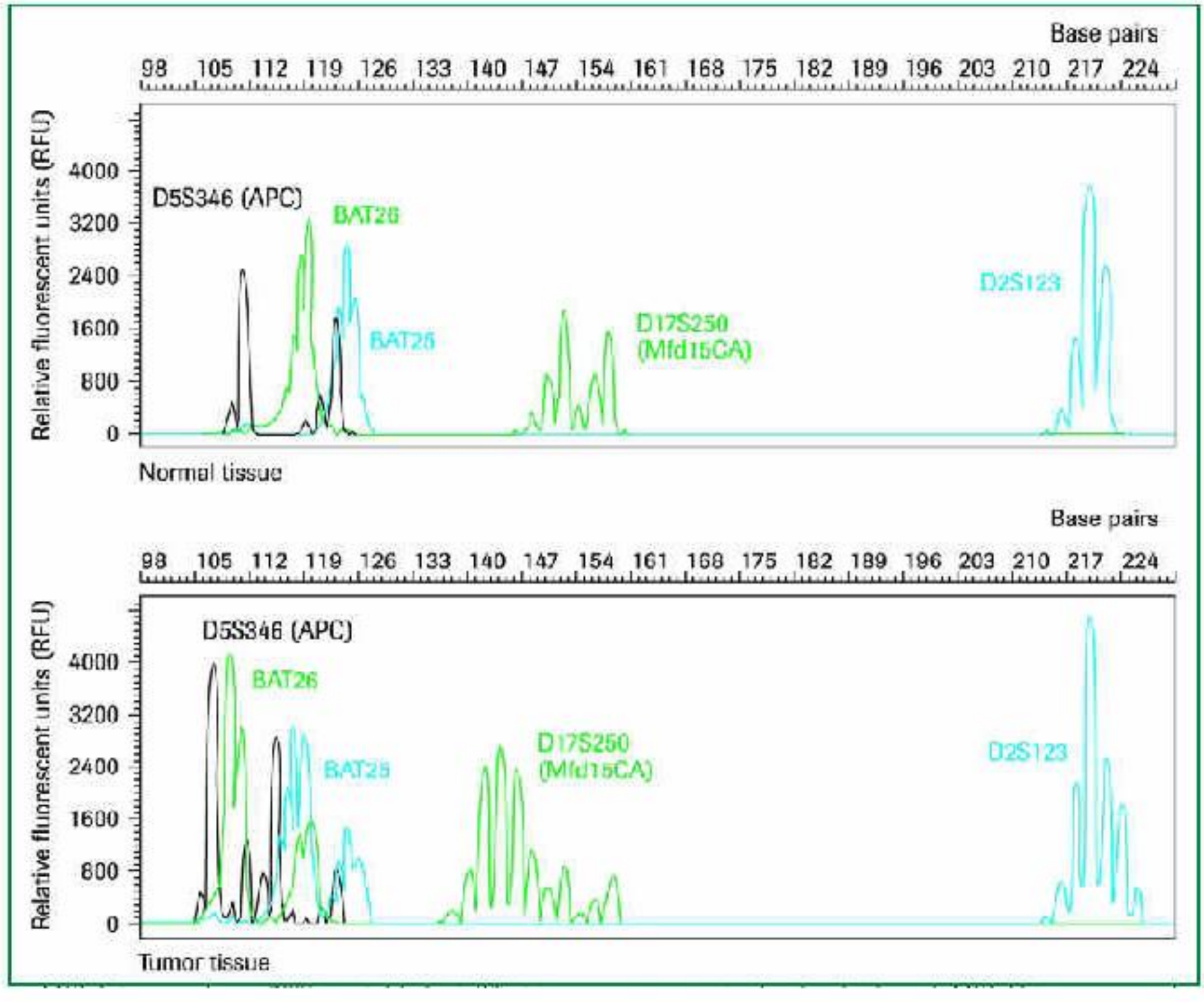
Disease	Gene /Locus	Protein	Schematic representation (not to scale)	Proposed Mechanism of Disease
Fragile X syndrome	FMR1 (FRAXA) Xq27.3	FMR-1 protein (FMRP)		Loss of function: Loss of FMR1: abnormal RNA metabolism
Fragile XE syndrome	FMR2 (FRAXE) Xq28	FMR-2 protein		Loss of function: Loss of FMR2: disruption of neuronal gene regulation?
Friedreich ataxia	X25 9q13-21.1	Frataxin		Loss of function: Reduced frataxin in mitochondria causing altered iron homeostasis and mitochondrial dysfunction
Myotonic Dystrophy	DMPK 19q13	Myotonic dystrophy protein kinase (DMPK)		Loss and/or gain of function: Reduced DMPK: disruption in kinase activity Cis-effects: silencing in the DM region Dominant effects on RNA processing (CUG-binding proteins)
Spinocerebellar ataxia type 8	SCA8 13q21	None		Loss of function? Abnormal RNA (antisense) regulation?
Spinocerebellar ataxia type 12	SCA12 5q31-33	PP2A-PR55β		Loss of function? Disruption in phosphatase activity?

Hum Mol Genet 2000

Pattern of amplified PCR fragments from a colorectal tumor with high microsatellite instability

Genomic DNA from the tumor and normal tissue were amplified with the HNPCC Microsatellite Instability Test. The amplified fragment patterns from both tissues were analyzed with an ABI PRISM 310 Genetic Analyzer

Result: The fragment pattern from the tumor differs from the normal pattern at all five microsatellite loci. Thus, the tumor is 100% unstable and is classified as an MSI-H .



Interspersed repetitive sequences

Retrotransposable elements (45% of our DNA!!!)

.

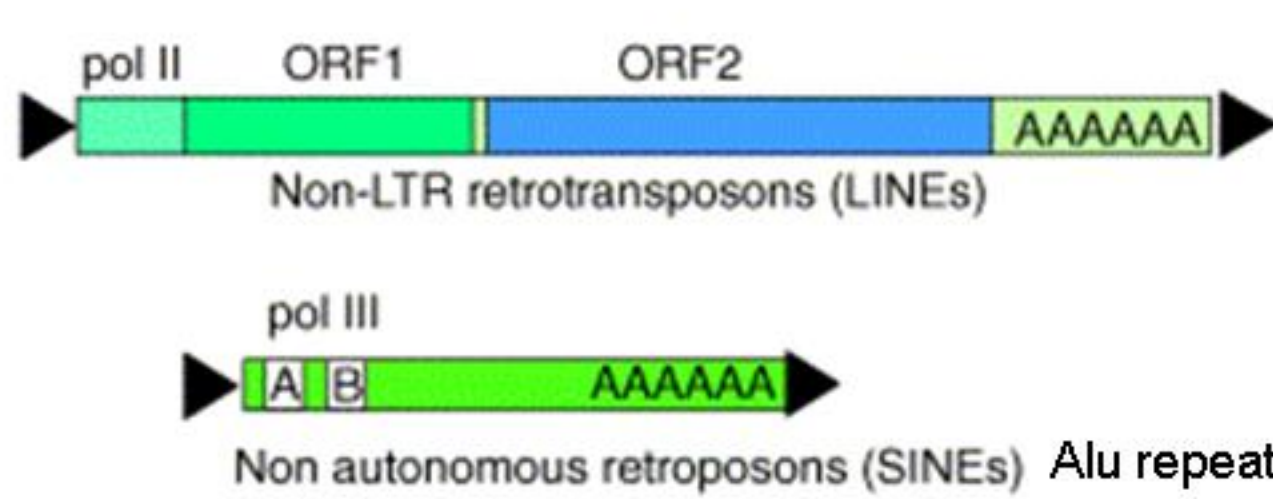
Polimorfizm typu *indel (insercja/delecja)*

Aktywacja w gonadach

Rola w indukcji crossing-over Sience 2005

Pośredniczą w rearanżacjach genomowych, szczególnie w indukowaniu strukturalnych aberracji chromosomowych

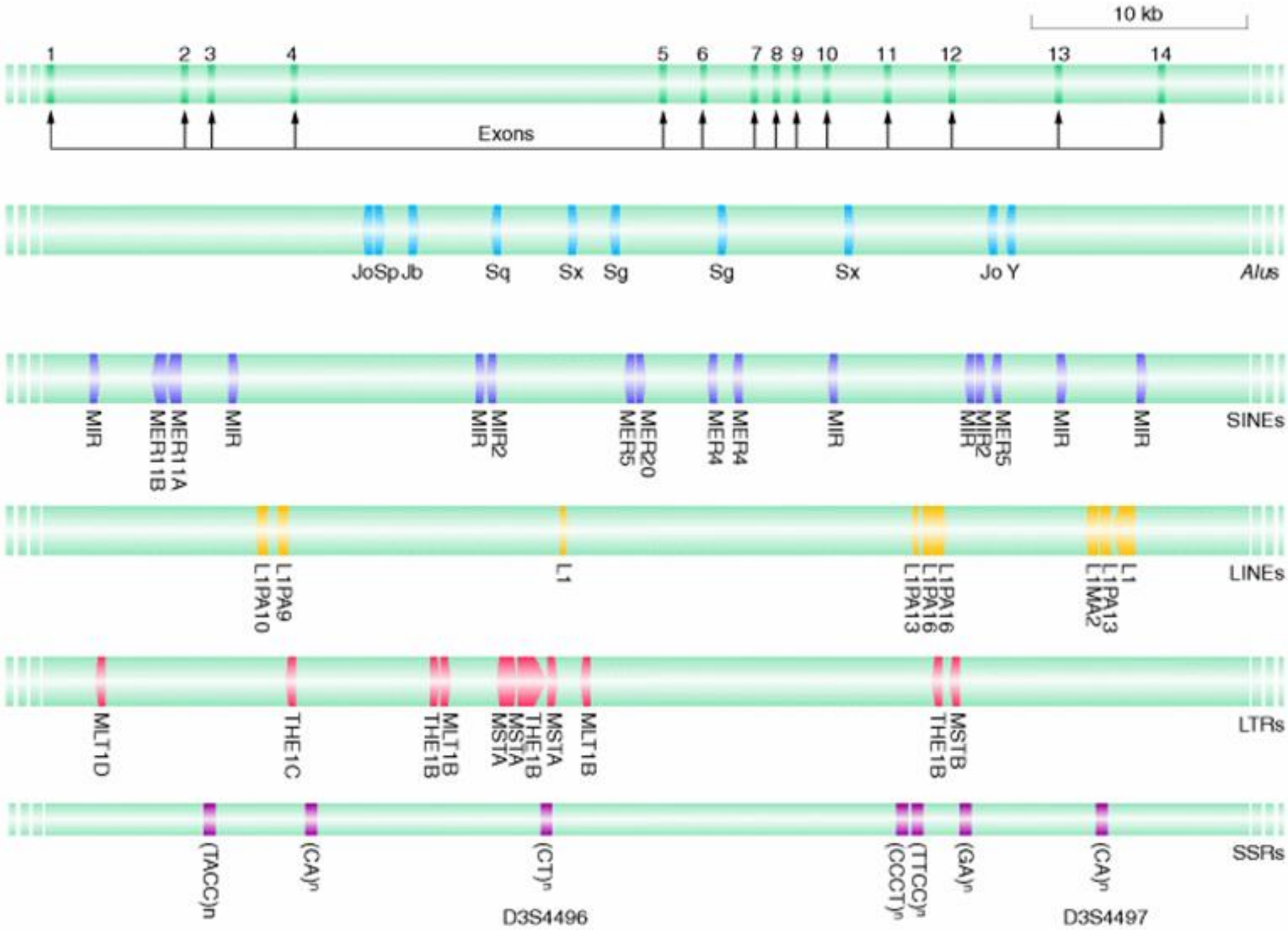
Sporadycznie (<1%) wywołują chorobotwórcze mutacje



Current Opinion in Genetics & Development

2003, 13:651

Repetitive elements in *HGO* gene (defective in alcaptonuria)



HLA – najbardziej polimorficzny układ genetyczny człowieka

Cząsteczki i geny HLA

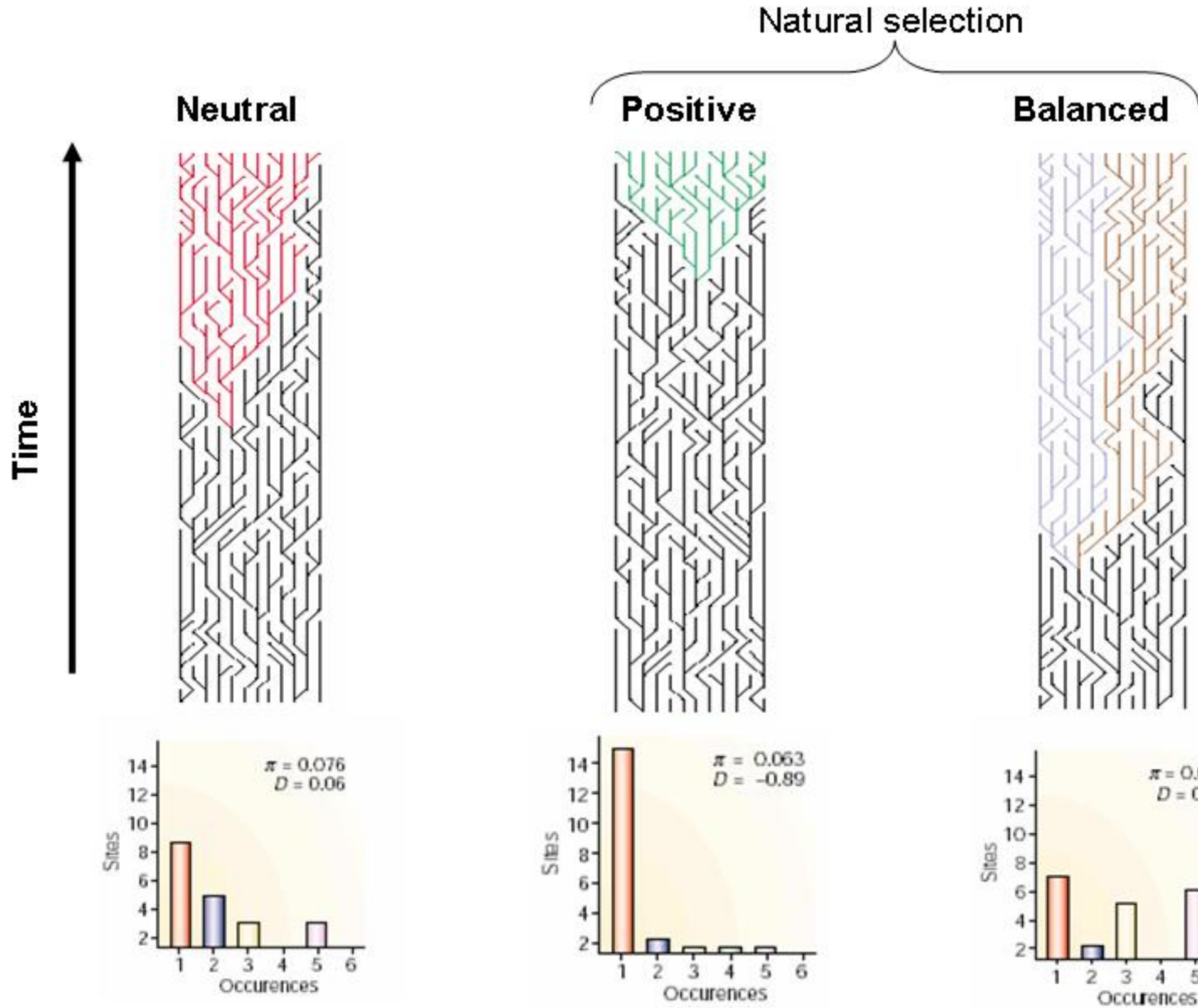
DP			DQ			DR					B	C	A
B1	A1		B1	A1		B1	B3	B4	B5	A	B	C	A
99	*22*		*53*	*22*		*396*	*30*	*10*	*15*	*3*	*499*	*125*	*258*

HLA-complex

Gene loci

Alleles (2002)

Polimorfizm HLA jest skutkiem selekcji równoważącej (*balanced selection*)

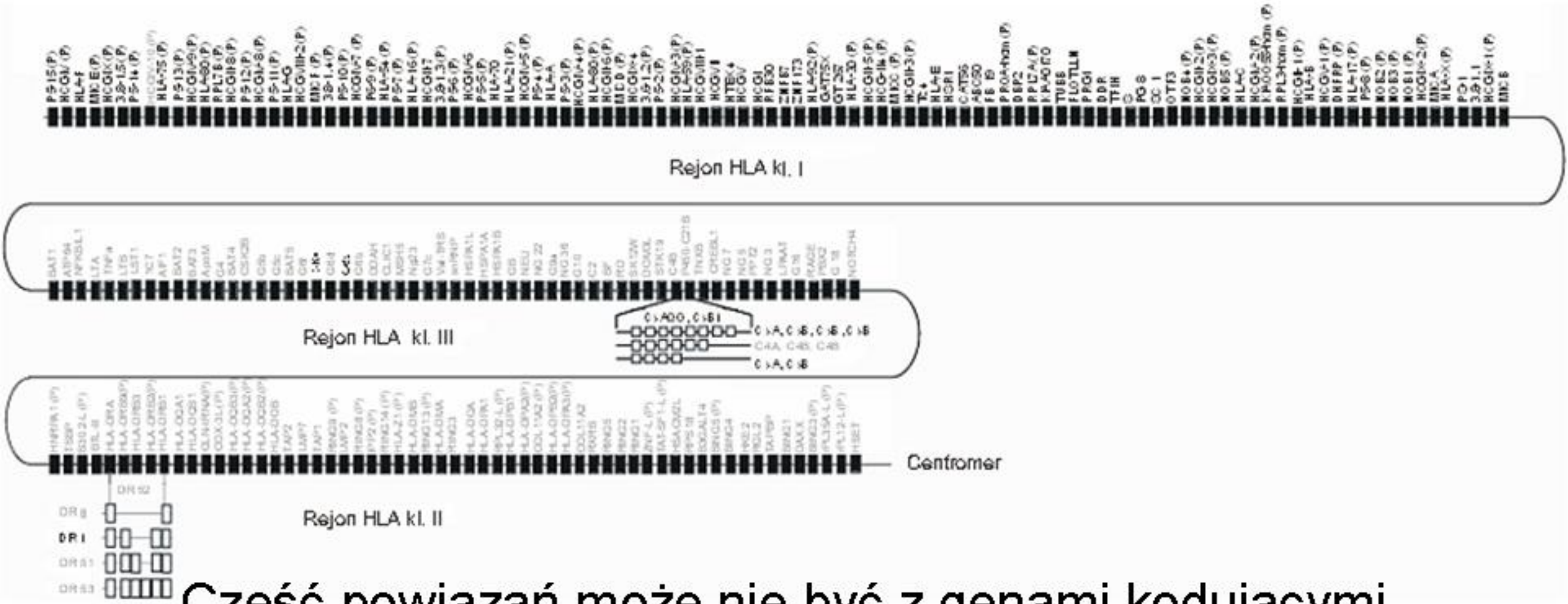


Powiązania HLA z chorobami

- Zzsk (zesztywniające zapalenie stawów kręgosłupa) - B27
- Celiaklia - DQ2 (DQA1*0501, DQB1*0201)
- Narkolepsja – DRB1*15 (DR2)

- Kilkadziesiąt innych bez znaczenia w klinice

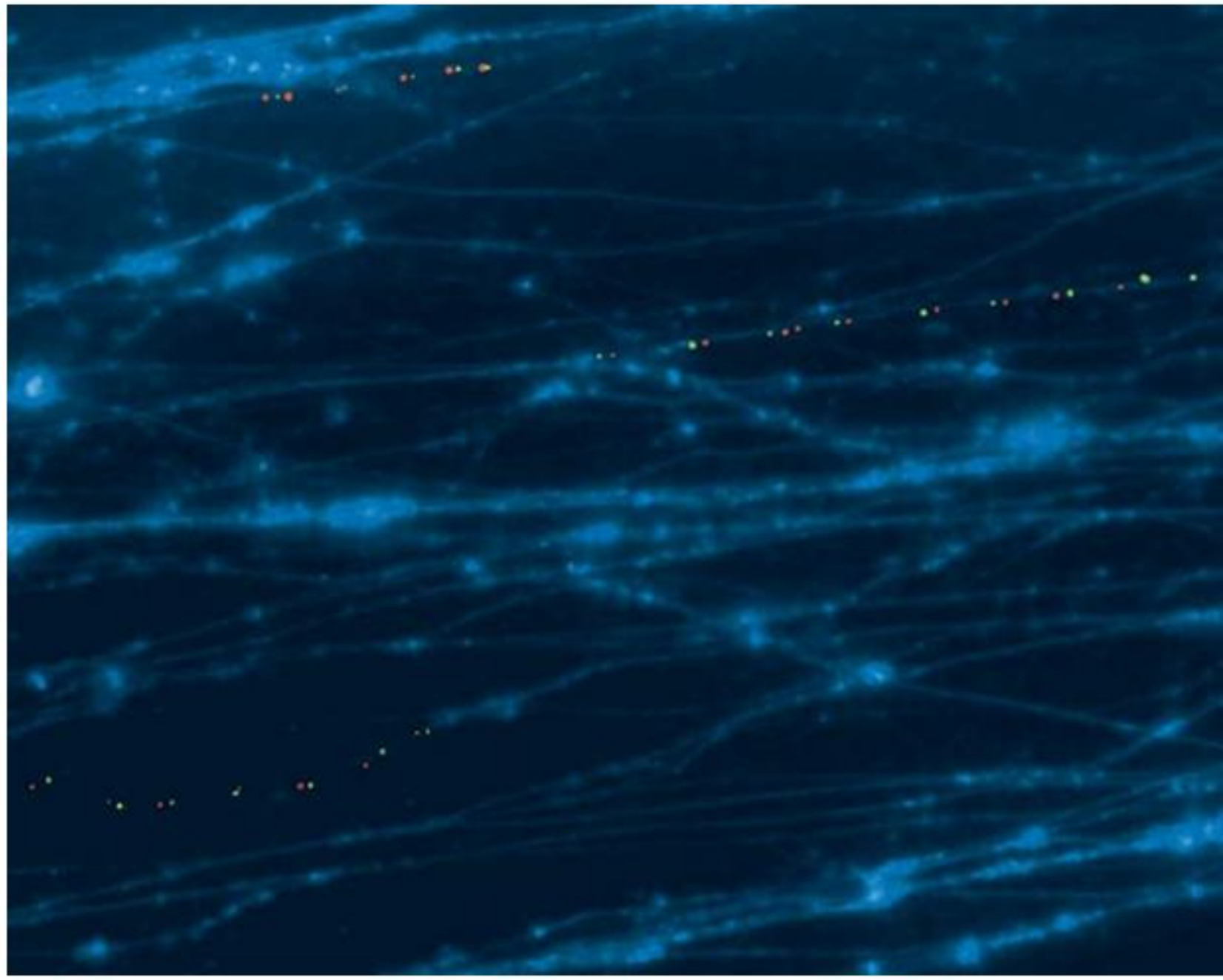
Układ HLA zawiera dużo innych genów (u człowieka region o **największej** gęstości genów)



Część powiązań może nie być z genami kodującymi cząsteczki HLA

(np. hemochromatoza, zawał serca?)

Segmental duplications / copy number polymorphism



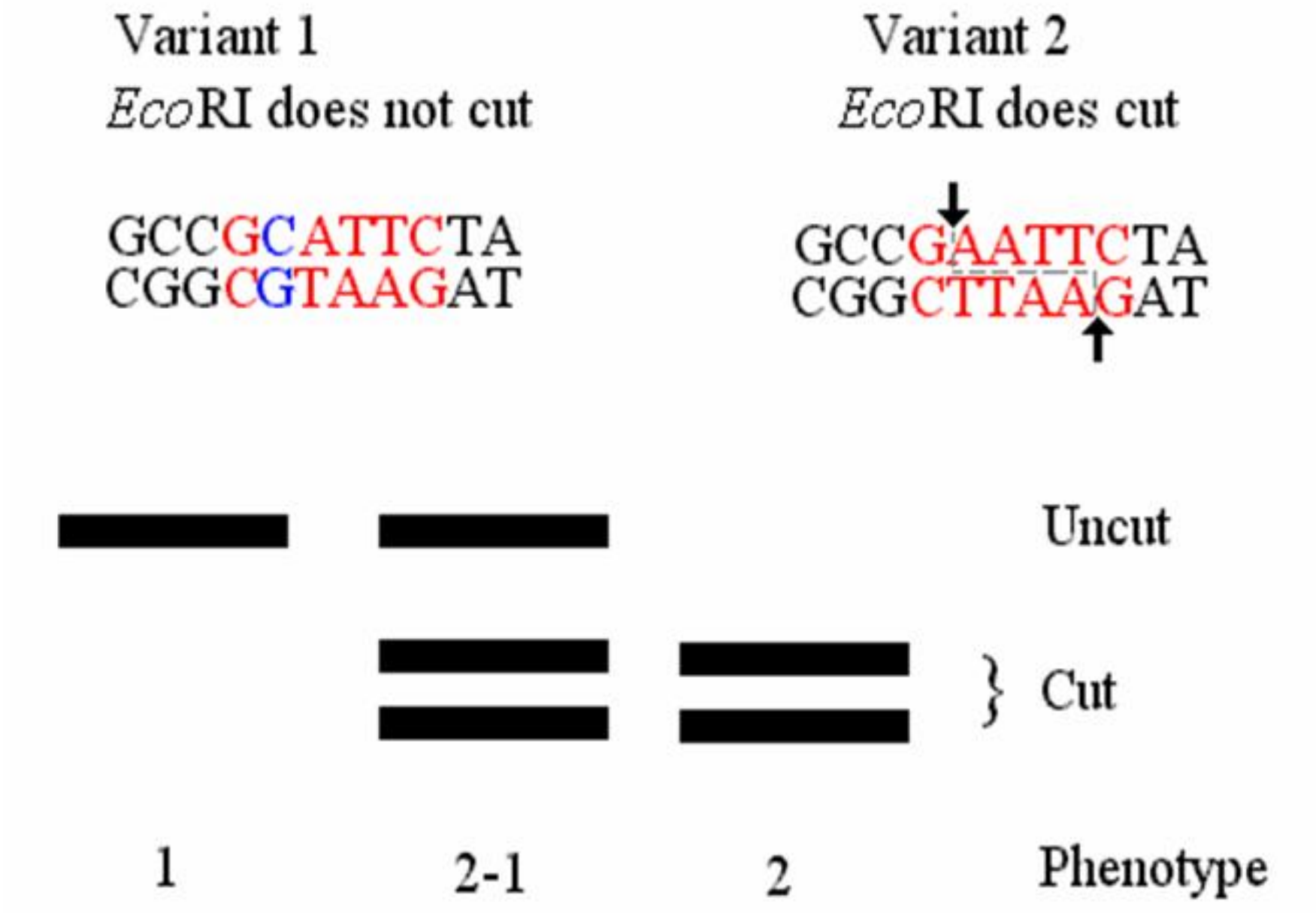
Hybridization of a 5' amylase gene probe (red) and a 3' amylase gene probe (green) to DNA fibers (blue) from three different individuals, each with a different number of tandem copies of the variable segment.

Nature 2004

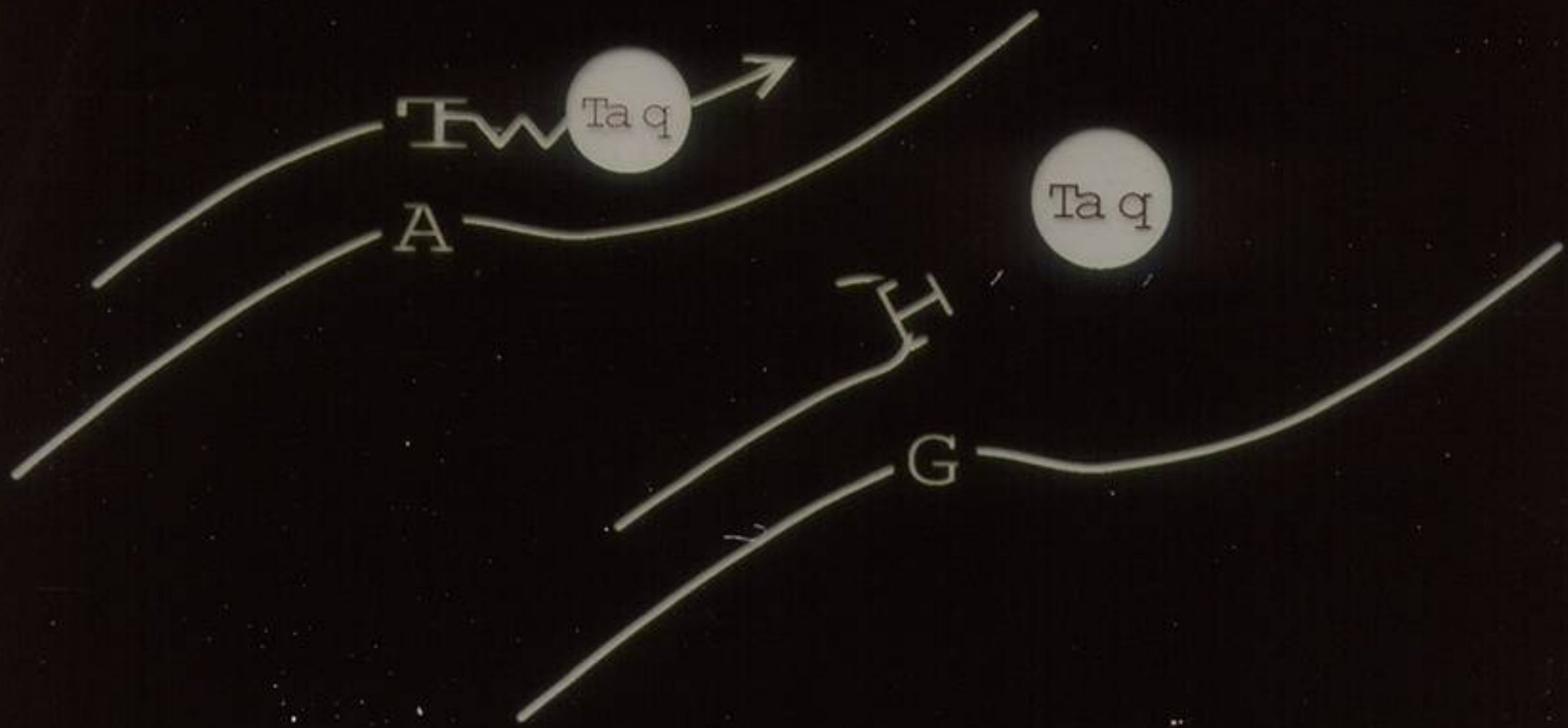
Metodyka analizy polimorfizmu DNA

Machine for PCR - thermocycler

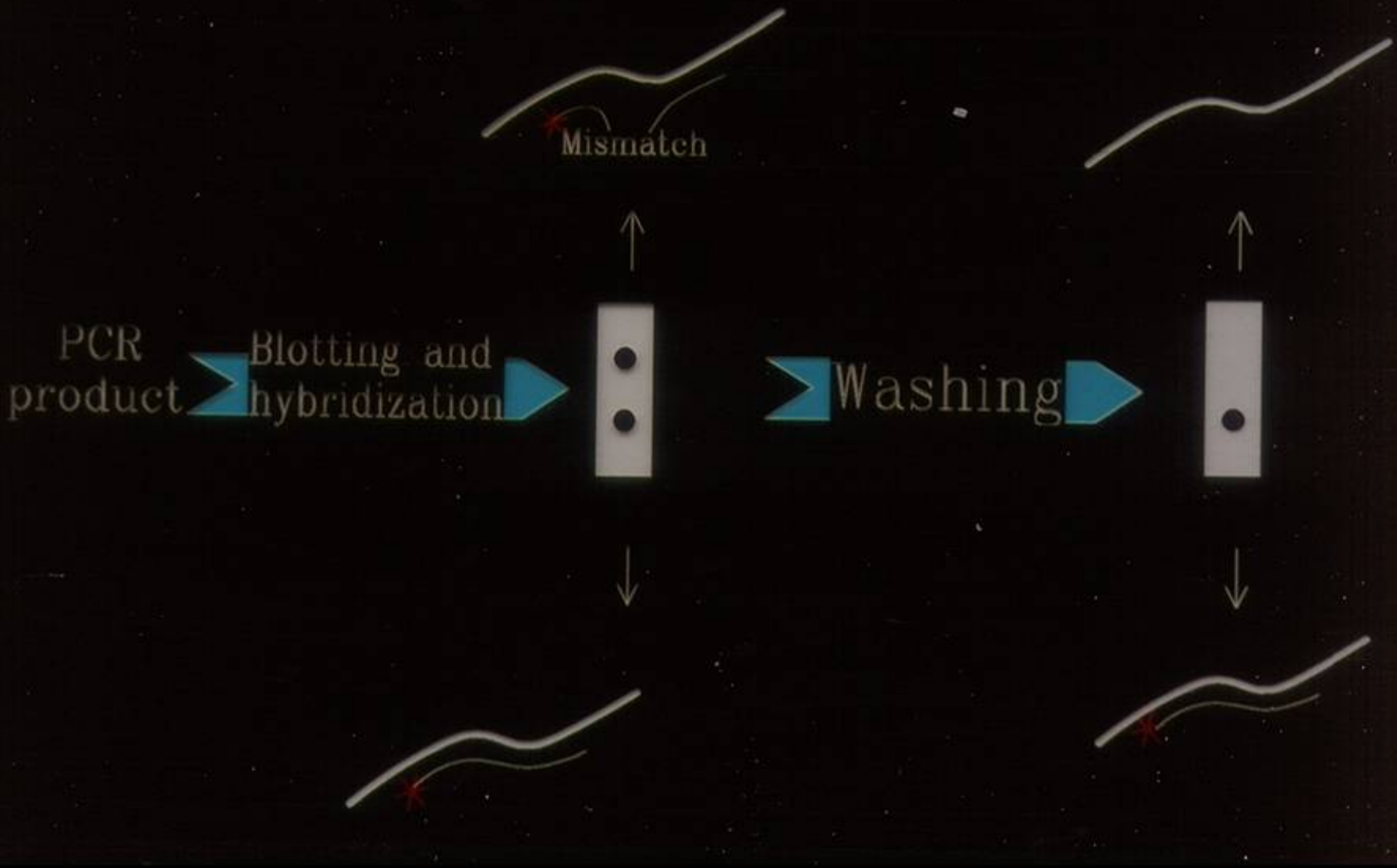
PCR-RFLP


Variant 1
*Eco*RI does not cut
GCCGCATTCTA
CGGCGTAAGAT
Variant 2
*Eco*RI does cut
GCCGAATTCTA
CGGCTTAAGAT
Uncut
Cut
1
2-1
2
Phenotype

Allele-specific PCR (AS-PCR)

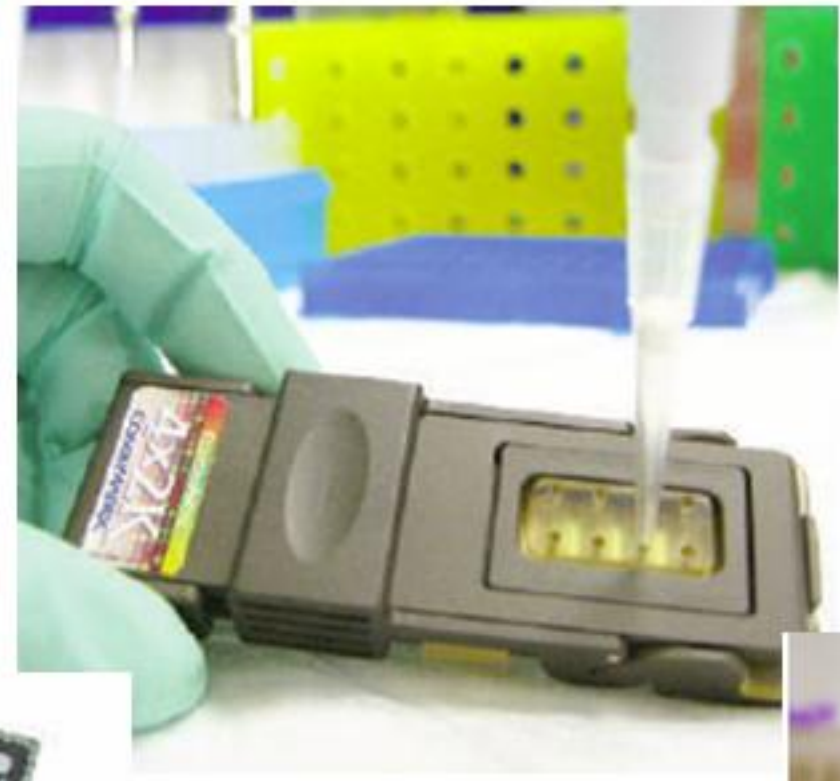


PCR-SSO analysis



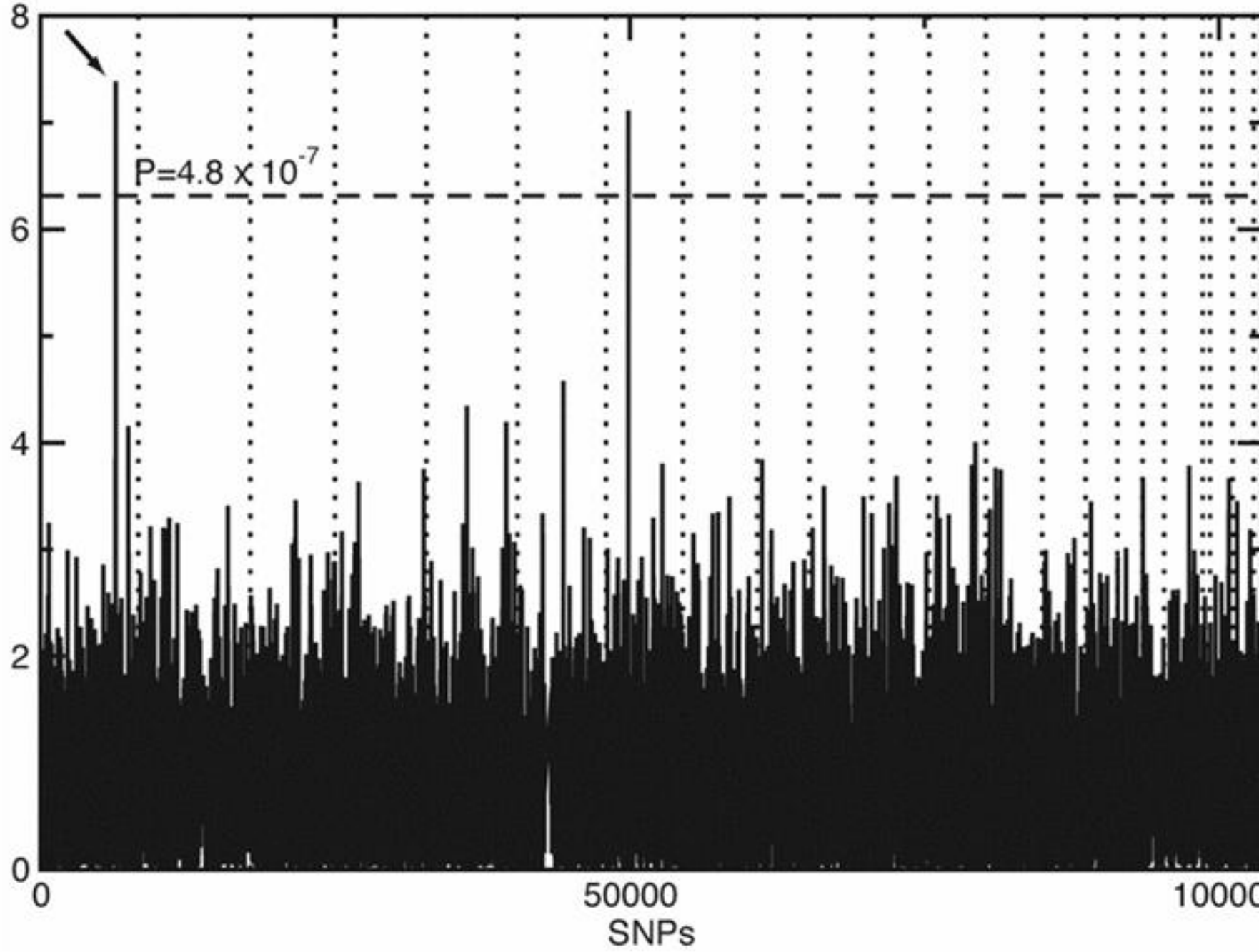
Micro-array and ‘gene chip’ technology

NATURE 437, 2005



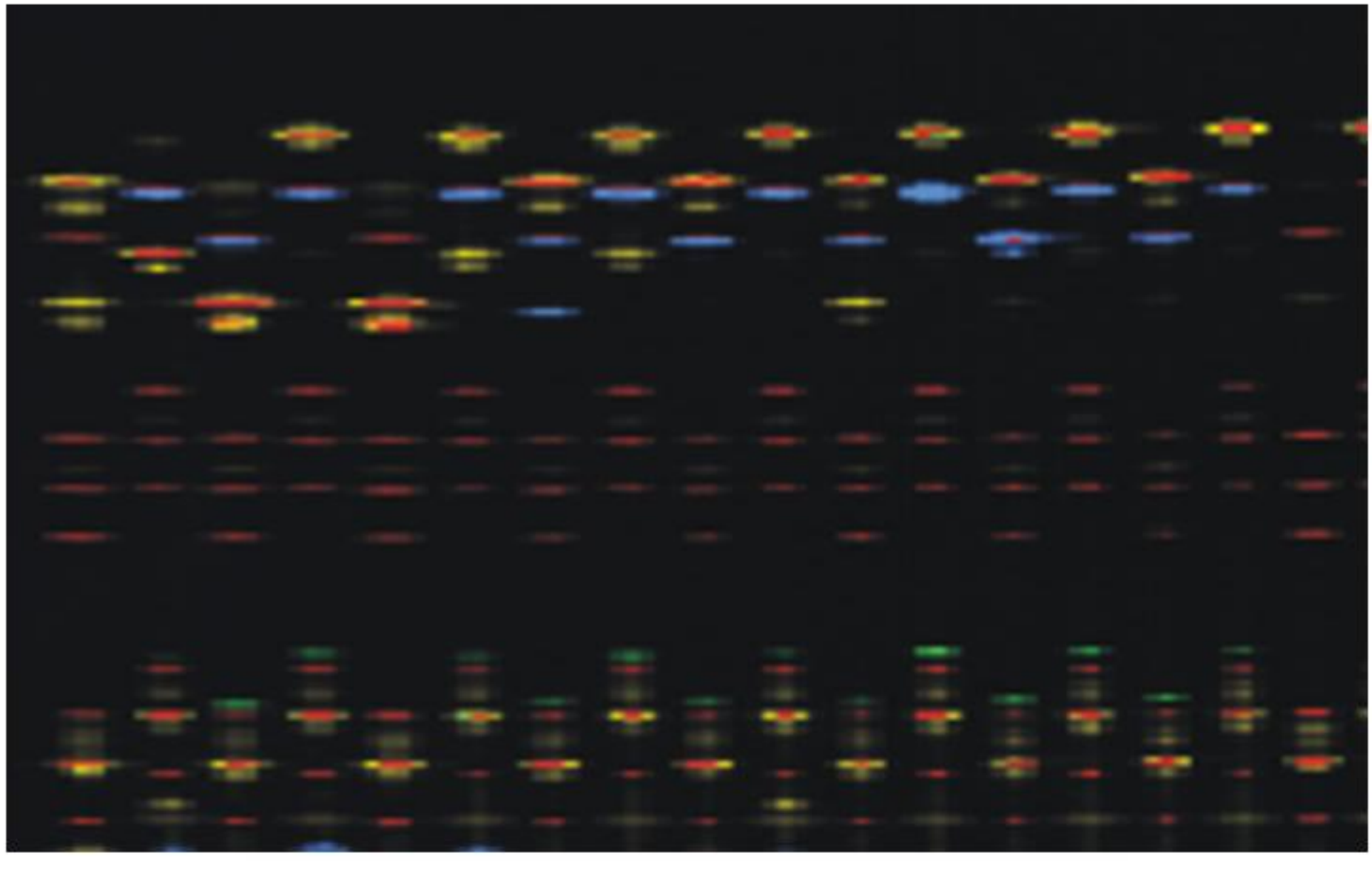
Whole-genome case-control association study for genes involved in AMD

P values of genome-wide association scan for genes that affect the risk of developing AMD. -log10(*p*) is plotted for each SNP in chromosomal order. The spacing between SNPs on the plot is uniform and does not reflect distances between SNPs on the chromosomes. The dotted horizontal line shows the cutoff for *P* = 0.05 after Bonferroni correction. The vertical dotted lines show chromosomal boundaries. The arrow indicates the peak for SNP rs380390, the most significant association, which was studied further

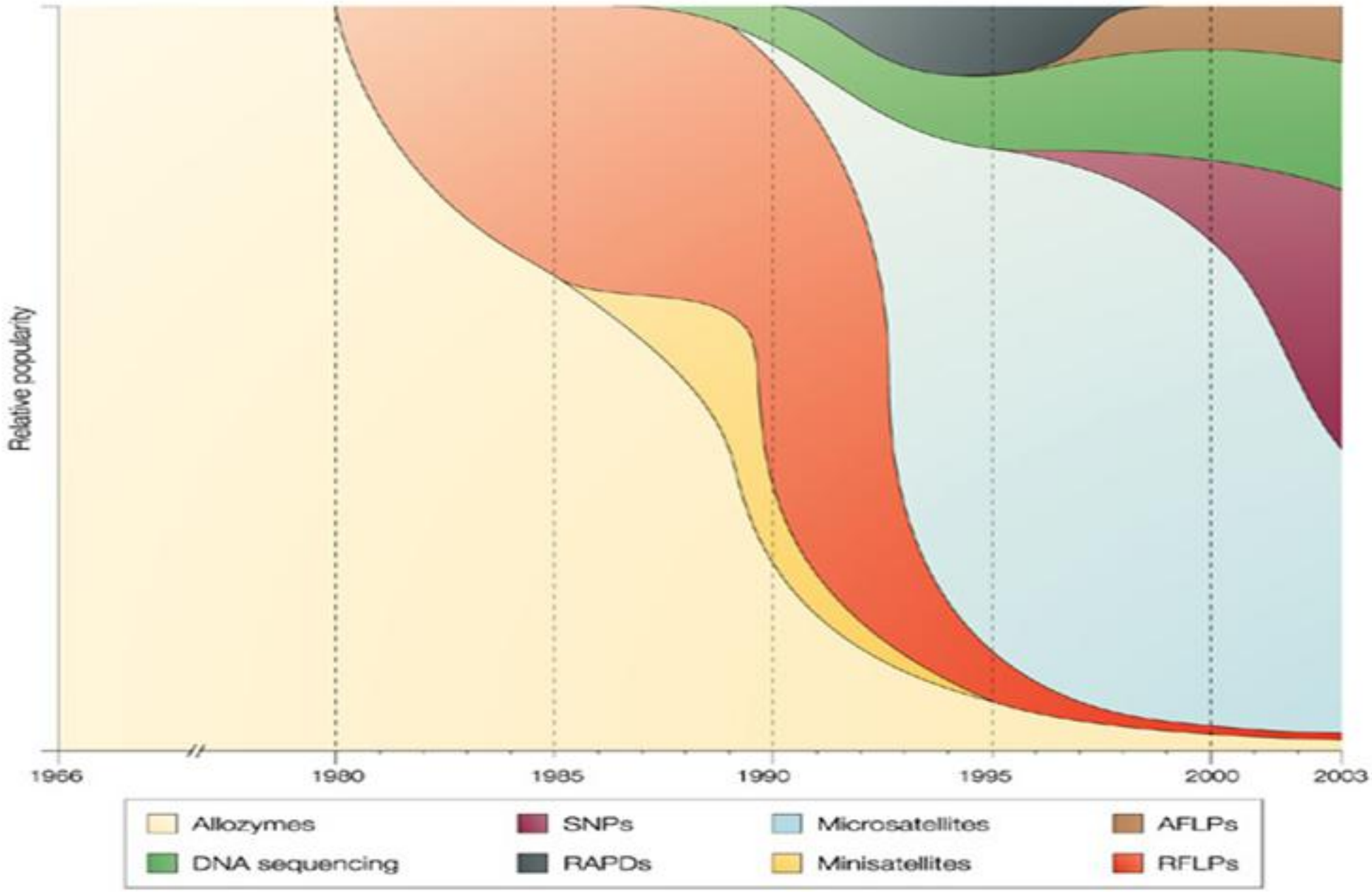


pdf Science 2005

Electrophoretograms of PCR products from a STR multiplex



Polymorphic genetic markers



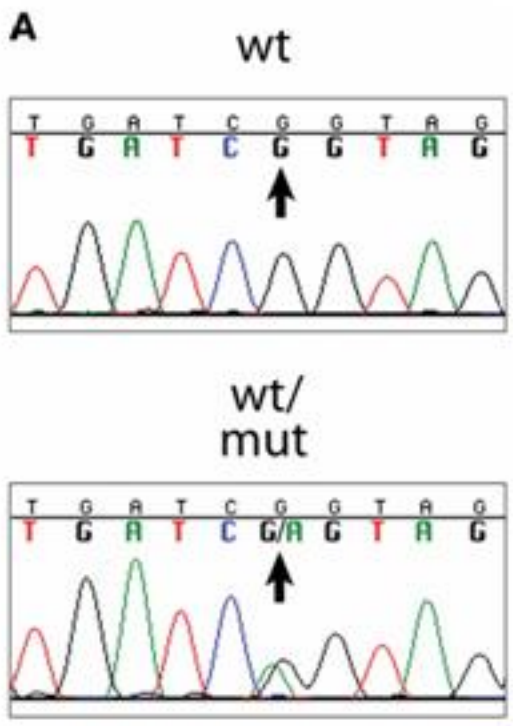
AFLP, amplified fragment length polymorphism; RAPD, randomly amplified polymorphic DNA; RFLP, restriction fragment length polymorphism; SNP, single nucleotide polymorphism.

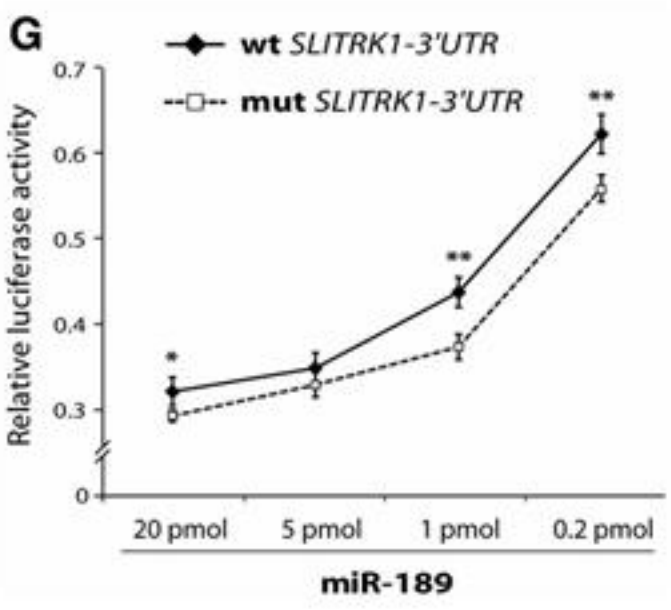
Nature Rev Gen 5 63-69 (2004)

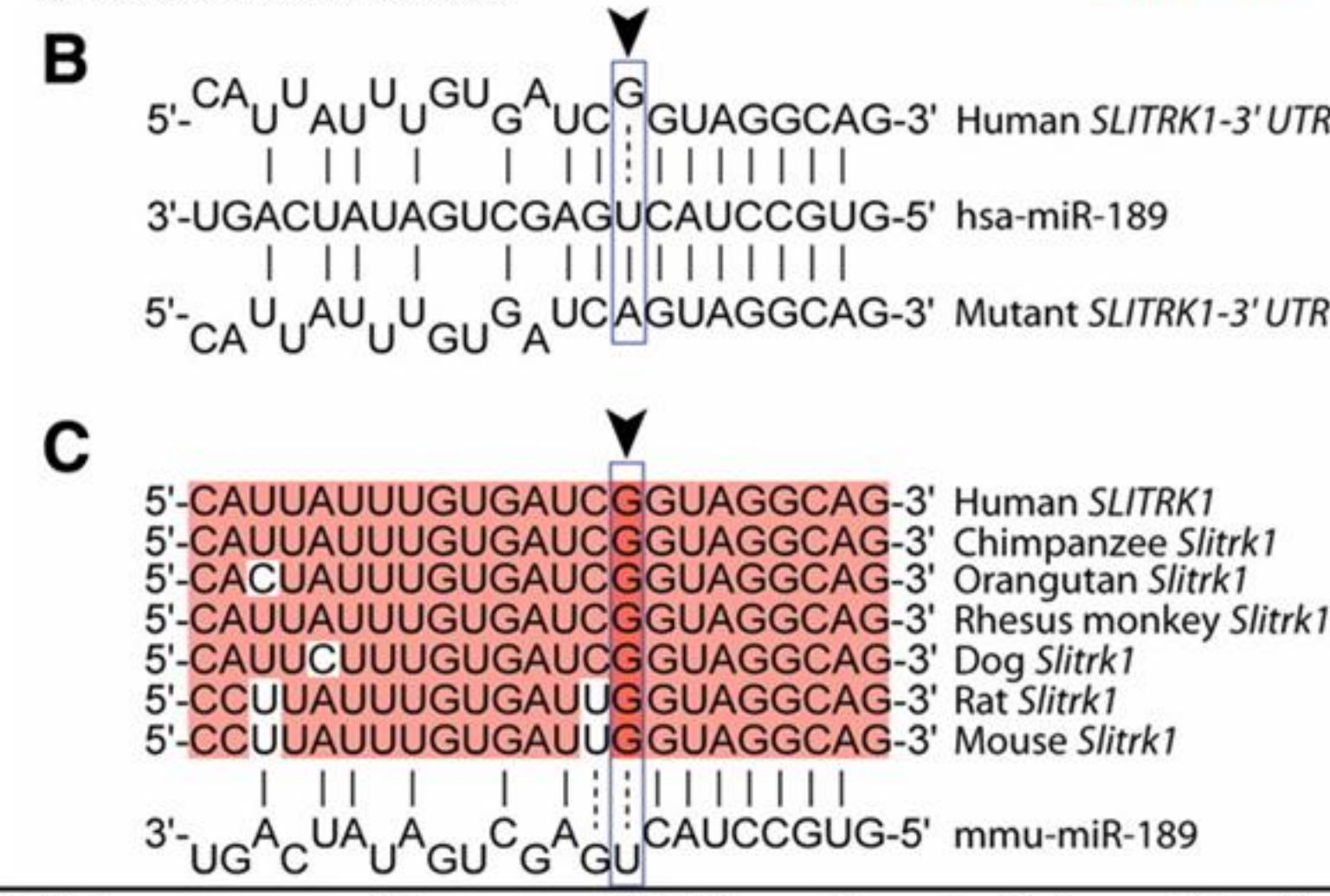
Dziękuję !

microRNA binding site mutation in 3' UTR of SLITRK1 may predispose toTourette's Syndrome

Science 310: 317

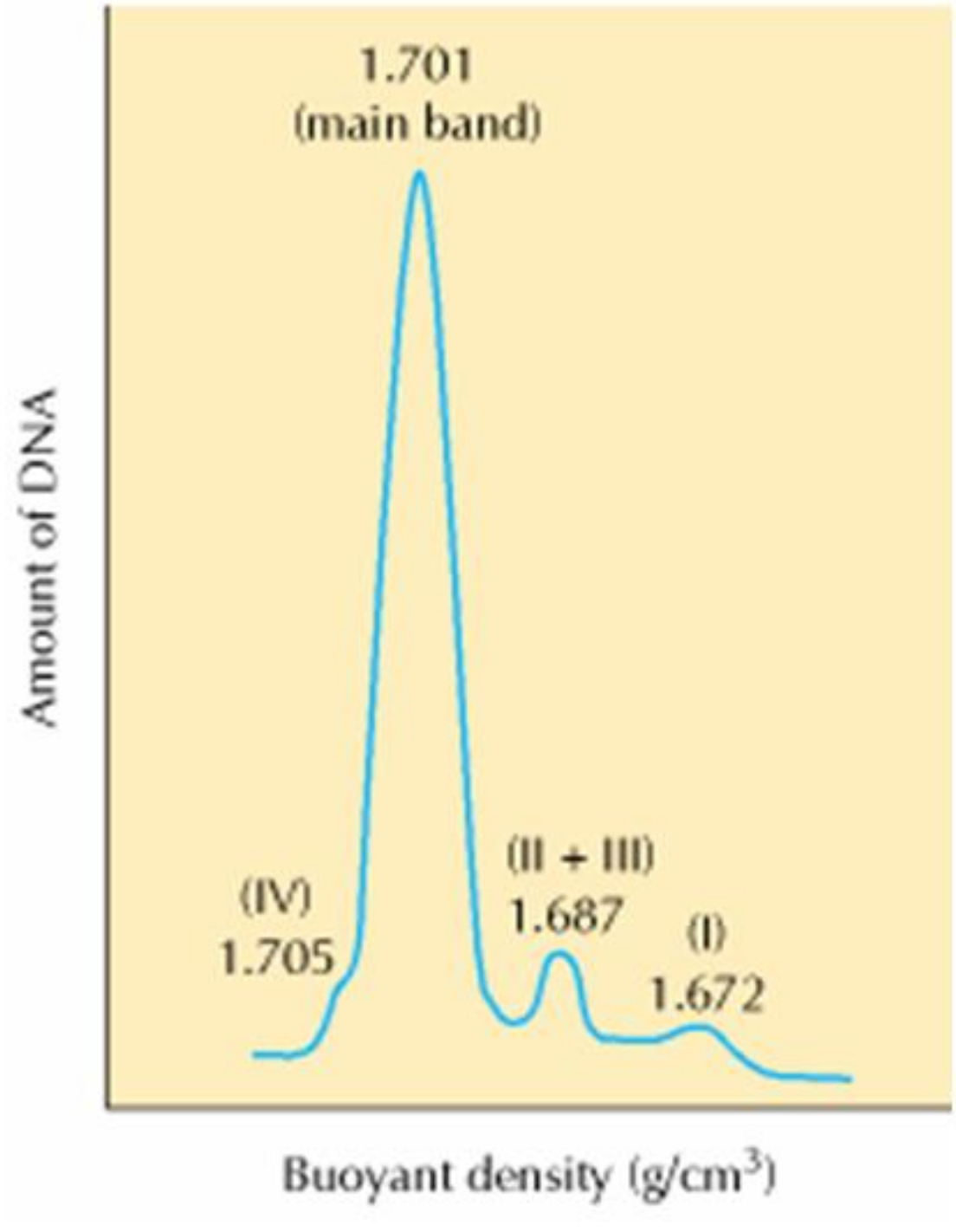




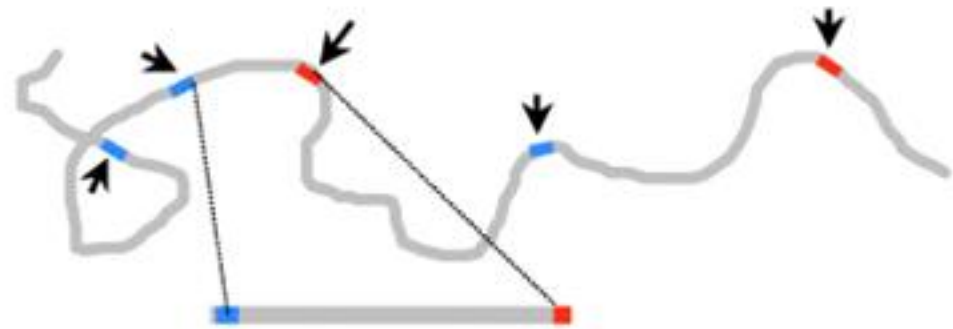


(**A**) The sequence of the normal 3'UTR and the substitution of G to A found in patients. (**B**) The substitution maps within a predicted miRNA binding site for miR-189. Base pairing is indicated by a solid (Watson-Crick) or a dashed (wobble) vertical line. (**C**) Conserved bases in the binding domain are shown in red. (**G**) Relative luciferase activity in the presence of miR-189 is shown for the wild-type (wt) *SLITRK1* 3'UTR (solid line) and mutant (mut) *SLITRK1* 3'UTR, containing the substitution of G to A

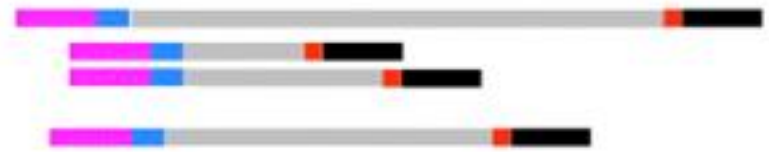
Centrifugation in a CsCl gradient separates satellite DNAs (designated I-IV) from the main band of genomic DNA



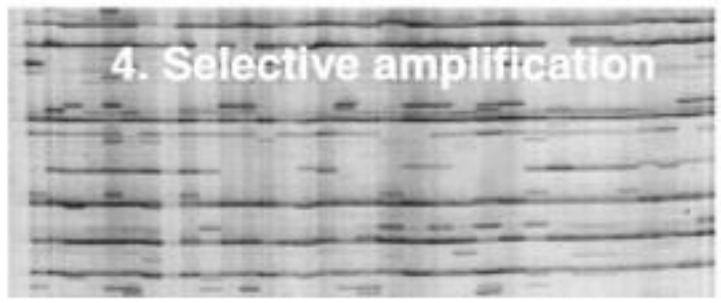
Principle of AFLP (amplified fragment length polymorphism)



1. Digestion of genomic DNA with two different restriction enzymes



2. Ligation of adaptors



DNA polymerase slippage as mechanism for STR mutations