A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC

Nature Genetics (38):1166-1172
20.10.2006

Tarmo Puurand

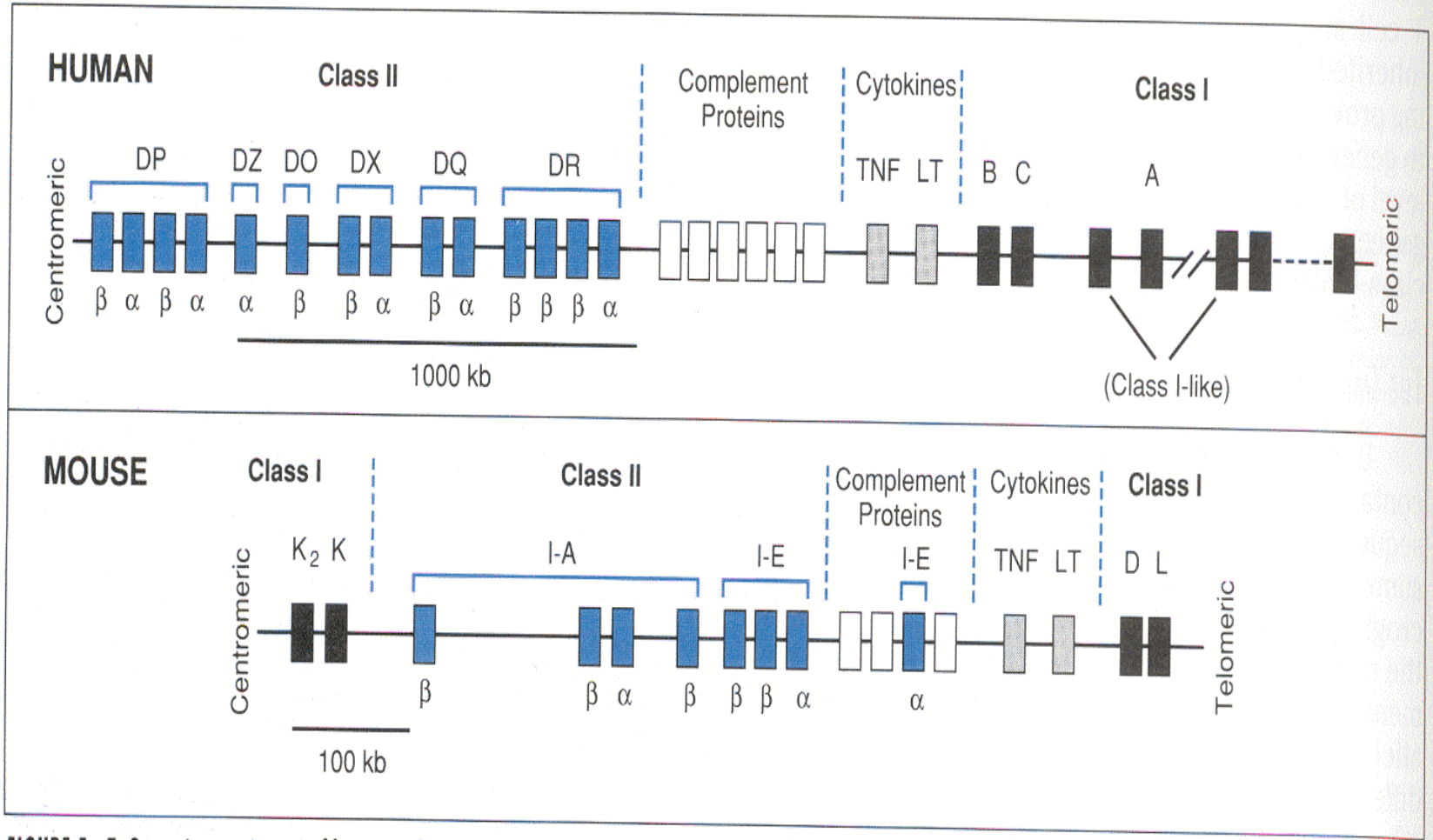


FIGURE 5–7. Genomic organization of human and mouse MHC loci. *Genes comprising each locus and intervening distances are of approximate relative size. The number of α and β genes in each human class II locus varies among alleles. The mouse cytokine genes (TNF and LT) are assigned tentatively based upon the location in humans. Note that each class Iα or class IIα or β gene consists of multiple exons that are not shown (see Fig. 5–8.). TNF, tumor necrosis factor; LT, lymphotoxin.*

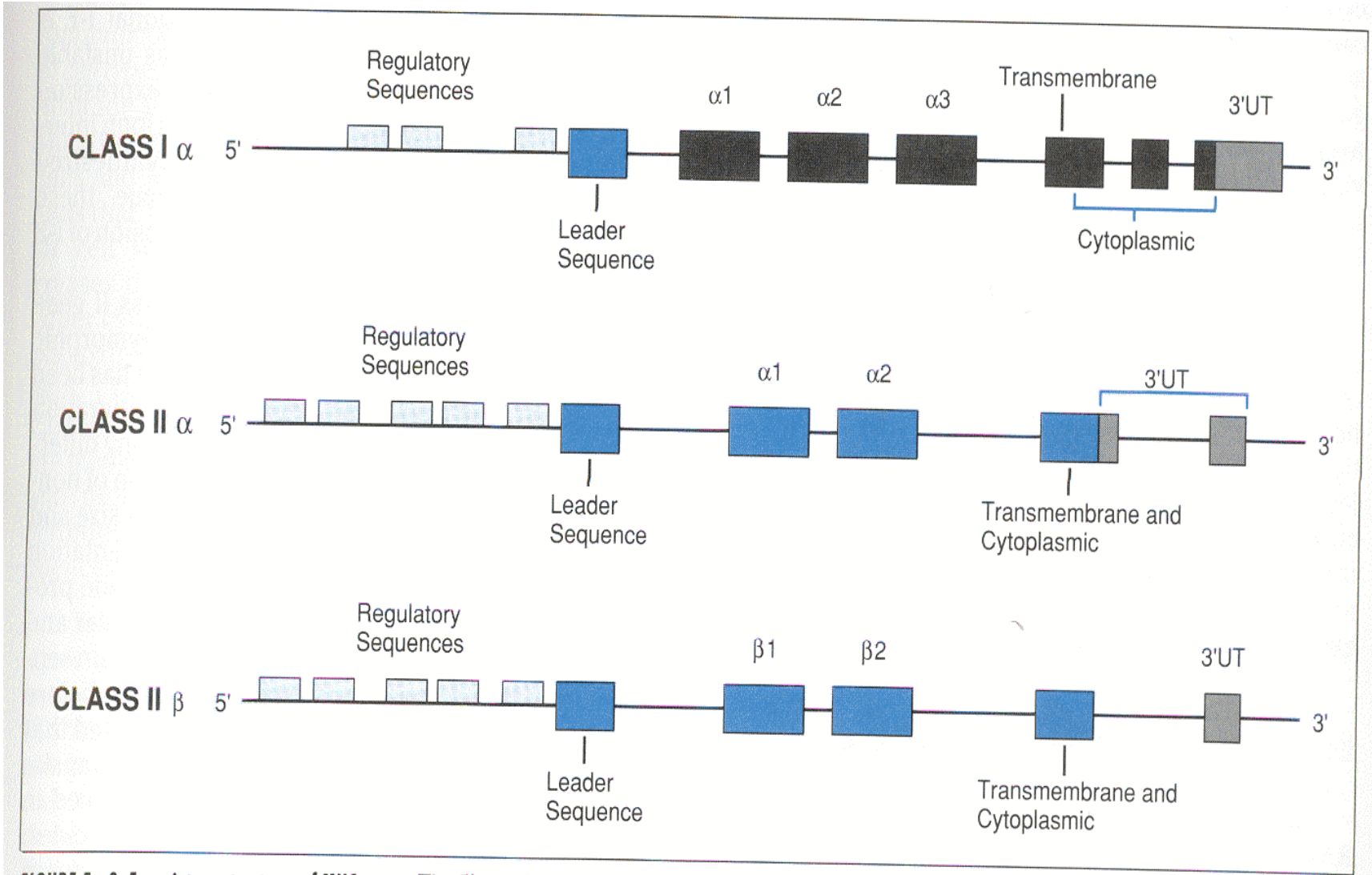


FIGURE 5 – 8. Exon-intron structures of MHC genes. *The 5′ regulatory sequences include promoter sequences, interferon-responsive elements, and multiple DNA sequences unique to class I or class II genes. 3′ UT indicates the 3′ untranslated sequence. Note that exons and introns are not shown to scale.*

Allele Information

HLA Class I Alleles:	904
HLA Class II Alleles:	620
HLA Alleles:	1524
Other Alleles:	64

Class I

A	B	C	E	F	G	H	J	K	L
258	499	125	6	1	15	0	0	0	0

Class II

DRA	DRB	DQA1	DQB1	DPA1	DPB1	DMA	DMB	DOA	DOB
3	396	22	53	20	100	4	6	8	8

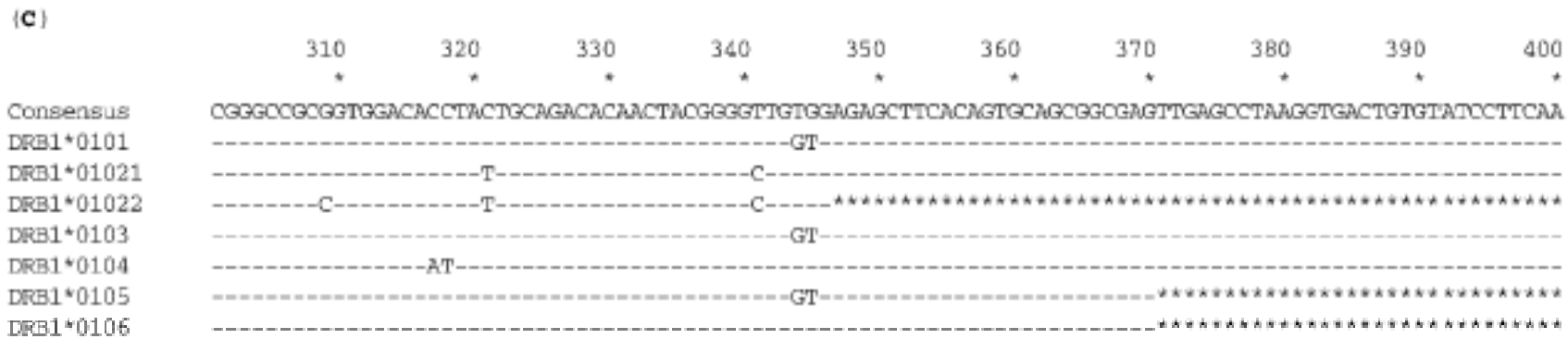
HLA-DRB

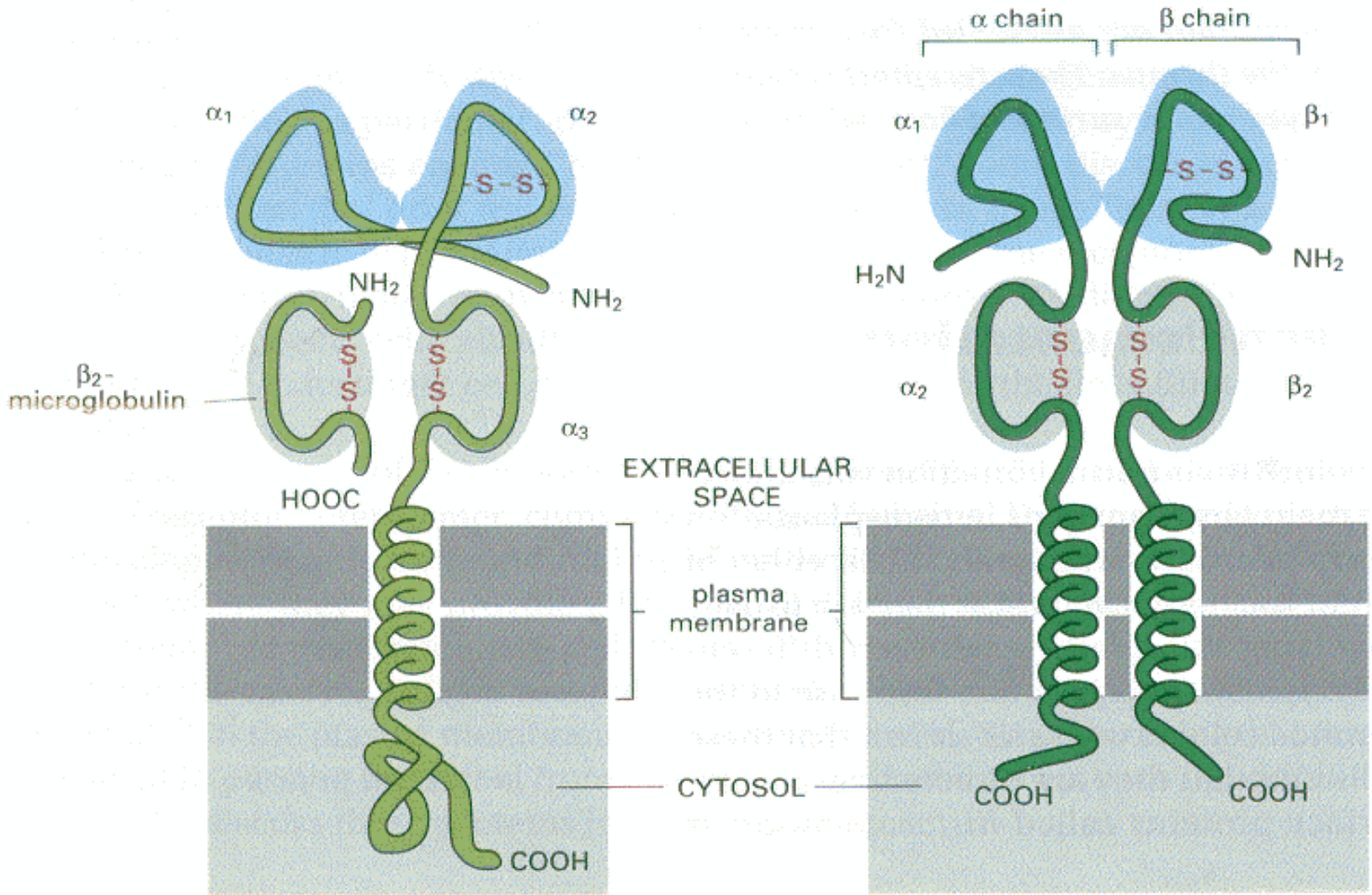
DRB1	DRB2	DRB3	DRB4	DRB5	DRB6	DRB7	DRB8	DRB9	TOTAL
321	1	38	12	17	3	2	1	1	396

Other non-HLA Genes

MICA	MICB	MICC	MICD	MICE	TAP1	TAP2	LMP2	LMP7
54	0	0	0	0	6	4	0	0

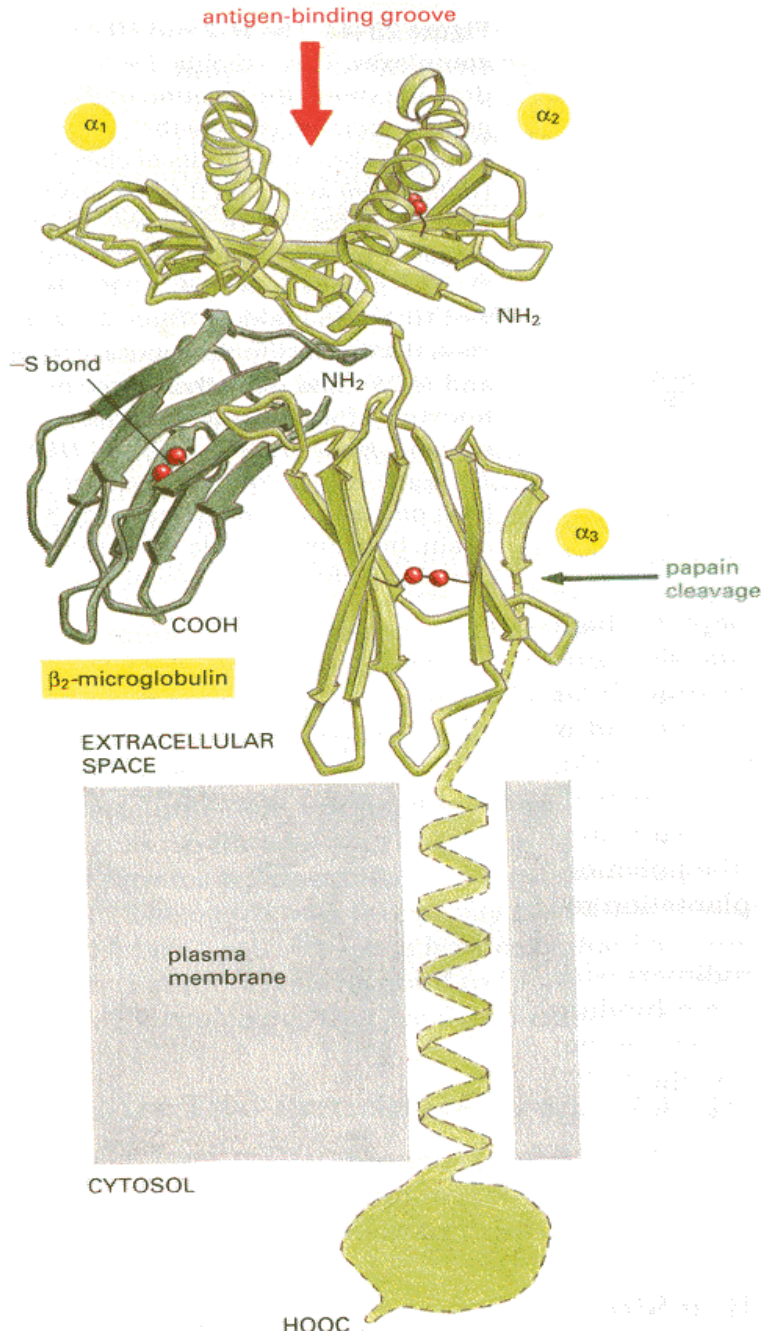
HLA geenide järjestused





(A) CLASS I MHC PROTEIN

(B) CLASS II MHC PROTEIN


antigen-binding groove
α1
α2
NH2
–S bond
NH2
α3
papain
cleavage
COOH
β2-microglobulin
EXTRACELLULAR
SPACE
plasma
membrane
CYTOSOL
HOOC

Supplementary Table 1. Partial Summary of Established HLA Associations, and Associations of Contemporary Interest

PHENOTYPE	HLA-A	HLA-C	HLA-B	C4	DRB1	DQA1	DQB1	DPB1	RR	Comments*
Autoimunity										
Acute anterior uveitis			*27							
Ankylosing spondylitis			*2702						~90	
			*2704							
			*2705							
Goodpastures syndrome					*1501/2				~16	
Graves disease			*0801		*0301				~4	
Hashimoto's thyroiditis					*0301	*0501				
Multiple sclerosis					*1501	*0102	*0602		~4	
Myasthenia gravis			*0801		*0301	*0501	*0201		~2.5	
Narcolepsy					*1501	*0102	*0602		~90	
Pemphigus vulgaris					*0402				~14	
Rheumatoid arthritis					*0101				~1-2	European descent
					*0401				~7	European descent
					*0404				~8	European descent
					*0405				~8	Oriental
					*1402				~2	Native American
SLE	*0101		*0801	C4AQ*0	*0301	*0501	*0201		~6	
Sjogren's syndrome			*0801		*0301				~3-20	DR2, DR5, DR11 some populations
Type 1 diabetes					*04	*0301	*0302		%	heterozygote
					*03	*0501	*0201			
					*09	*0301	*0303		%	heterozygote
					*03	*0501	*0201			
					*08	*0301	*0402		%	heterozygote
					*04	*0301	*0302			
					*0401	*0301	*0302		%	
					*0402	*	*		%	
					*0403	*	*		%	
					*0404	*	*		%	
Polymyositis/Dermatomyositis										MICA5.1/TNF2/TNFa2/DRB1*03 haplotype (8.1 ancestral haplotype)
										7.1 ancestral haplotype protective
Other immune disorders										
Behcet's disease			*5101 *57							
Bird-shot retinopathy	*A29		*12						49.9-224	
Celiac disease			*0801		*0301	*0501	*02			DQA1*1501/DQB102 in cis or trans
Crohn's disease										DR7 (not confirmed in all studies)
Ig deficiency										class II or class III?
Psoriasis		*0602								PSORS1 - cluster of loci near HLA-C
Sarcoidosis					*12/14				~8	BTNL2 near DRA1 implicated
			*07							various class I and II

Lofgren's					*0301		*0201		~21	
Mixed Cryoglobulinemia (HCV-associated)					DR11					DR11 phenotype is associated with a significantly increased risk for the development of type II MC in patients with chronic HCV infection. In contrast, HLA-DR7 appears to protect against the production of type II MC.
Hypersensitivity										
abacavir			*5701							mapped to class III HSP70?
allopurinol			*5801							
asthma										various class II - not seen in all cases - dependent on allergen. HLA-G associated in recent study
Berylliosis								*0201		DPB1 Glu69 susceptible/Lys69 protective
carbamazepine			*1502							
pigeon breeder's lung										class II
pollen-induced allergic rhinitis						*0302				chinese; 41 cases, 41 controls
Infections										
AIDS progression			*35Px							mostly HLA-B, protective: e.g. B*27/*57/*51
Leprosy/Tuberculosis						*02	*0503			Cambodian, 48 cases, 39 controls
Lyme disease					*0401					
Malaria			*5301							DRB1*1302 protective
SARS			*4601 *5401							
Kaposi's sarcoma - HIV associated					alleles with F13					DRB1 phe13 susceptible/gly13 protective
Other diseases										
Haemochromatosis	*03									due to linked HFE-class I gene
21-OH deficiency										CYP21 in class III region
cervical cancer										class II
nasopharyngeal carcinoma	*0207									chinese
smoking behavior				C4AQ*0						
hypertrophic cardiomyopathy			*51							A2-B51-DR2 haplotype, Asian Indian population (14 cases, 81 controls)
non-response to HBV vaccine					*07				relative odds [RO]=5.18	N=164
HCV - Sustained response to therapy			*44							
Gastric cancer					*04051					Japanese; 70 cases, 121 controls
N.B.										

There is a huge body of data on HLA and disease associations. A recent compilation describes over 200 separate conditions where MHC markers are associated with either disease susceptibility or severity of symptoms [Lechler, 2000]. In this Table are listed some of the more established associations, as well as others of contemporary interest. The earliest reference is given in most cases. Some of these early studies contained small patient numbers but generally have been confirmed in more than one population. For example, the association of SLE with various MHC haplotypes has been confirmed in over 66 studies on 25 populations but the precise etiologic locus or loci remain unidentified[Pich Books and reviews provide a more comprehensive literature [Tiwari, 1988][Lechler, 2000]. In many cases it is not possible to provide an accurate measure of the contribution of individual alleles to a condition. There are several explanations for this, including: 1. Multiple conflicting or underpowered studies. 2. Data from different populations. 3. Multiple susceptibility and protective alleles at different linked loci (e.g. HLA-DQ and -DR). 4. Dise This table summarizes some of the main associations as a starting point for accessing the literature.

HLA associated disiases

Table 1 Examples of HLA alleles associated with common disease and their tag SNPs

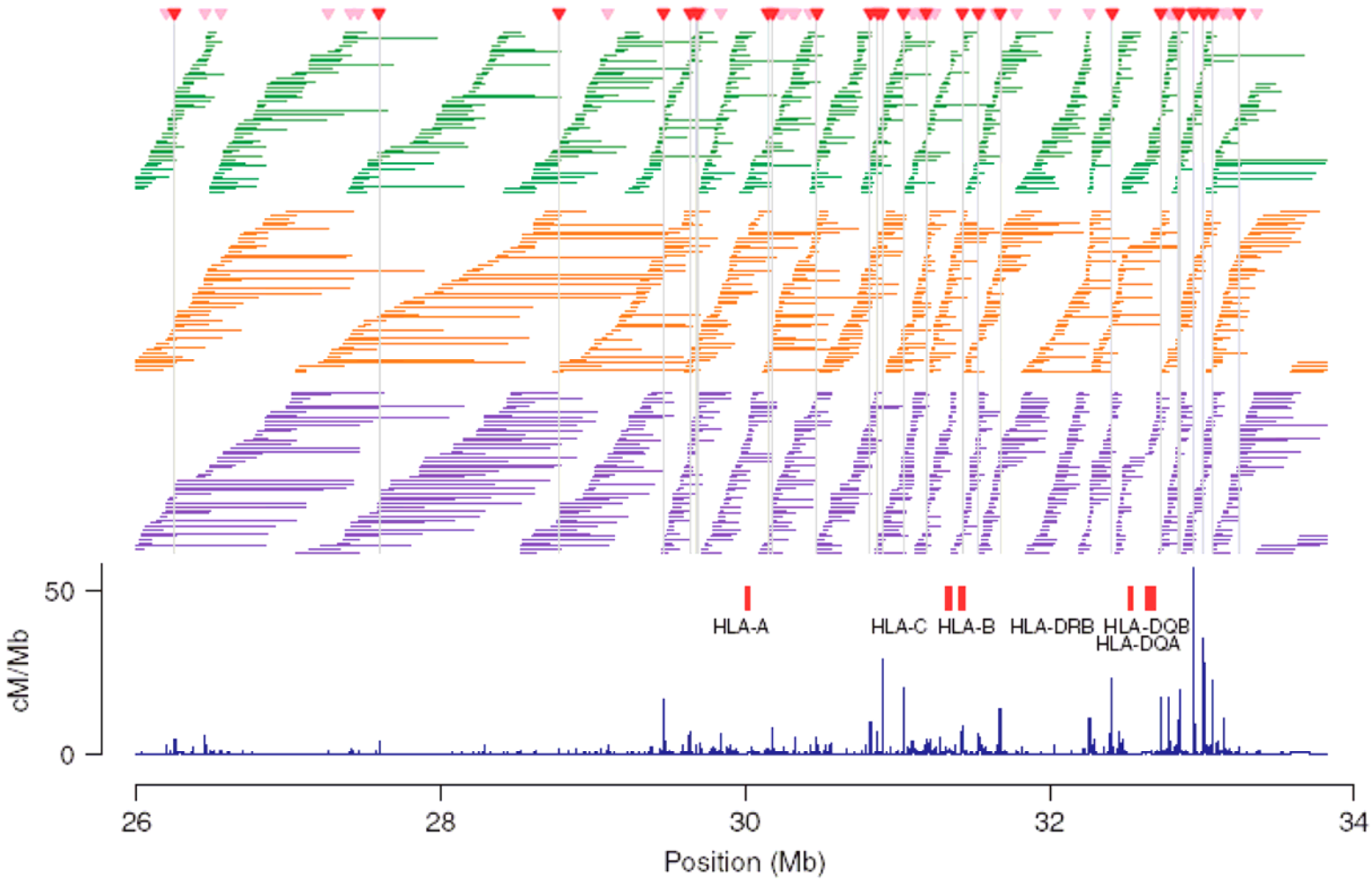
Phenotype[a]	Risk allele(s)	Estimated relative risk	Tag SNPs[b]	r^{2c}
Graves' disease or myasthenia gravis	B*0801/DRB1*0301/DQA1*0501/DQB1*0201	4 or 2.5	rs3129763(C) + rs4639334(C)	0.96
Multiple sclerosis	DRB1*1501/DQB1*0602	4	rs3135388	0.97
Multiple sclerosis	DQA1*0102	4	rs9268428(C) + rs6457594(A) + rs7451962(C)	0.90
Psoriasis	C*0602	5	rs887466(G) + rs4379333(C)	1.0
Celiac disease[d]	DQA1*0201/DQB1*0202 (DQ2.2)	1	rs4988889(T) + rs2858331(C)	1.0
	DQA1*0501/DQB1*0201 (DQ2.5)	7	rs4988889(T) + rs2858331(T)	0.93
SLE	DRB1*1501	2	rs3135388	0.97
Type 1 diabetes or SLE	DRB1*0301	4.5	rs2040410	0.87
Abacavir hypersensitivity	B*5701	4	rs2395029	1.0

[a]A more complete list of disease associated alleles can be found in **Supplementary Table 1**, and tags for the HLA alleles can be found in **Supplementary Table 3**. [b]Tags picked from CEU samples to capture the HLA risk allele. For multimarker (haplotype) tests, alleles of the individual SNPs are also listed in parentheses. For many HLA alleles, there are likely to exist multiple equivalent tags and predictors. [c]Coefficient of determination (r^2) between the identified HLA predictor and the HLA risk allele in the CEU panel. [d]DQA1*0201/DQB1*0202 (DQ2.2) has no effect on its own. Only when a person carries the DQ2.2 in combination with DQA1*0501/DQB1*0201 (DQ2.5) does it increase risk. The relative risk of DQ2.5 changes depending on the specific combination with other DQ types.

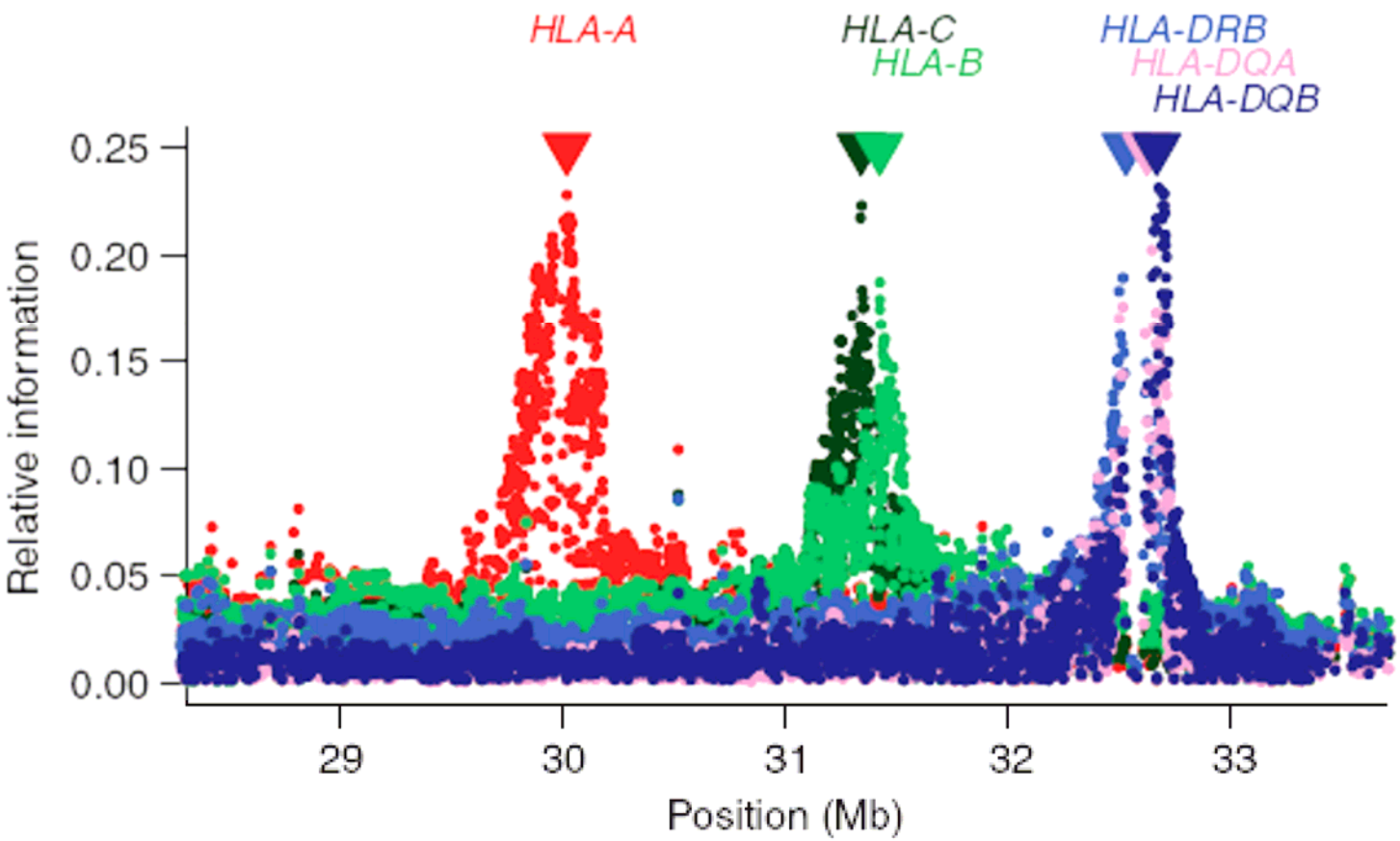
Study

- MHC region 7,5 Mb
- 361 individuals (SNP)
- 7543 SNPs and DIPs
- HLA typing 330 Dutch (DQA1, DQB1) and 332 UK trio (DRB1) samples

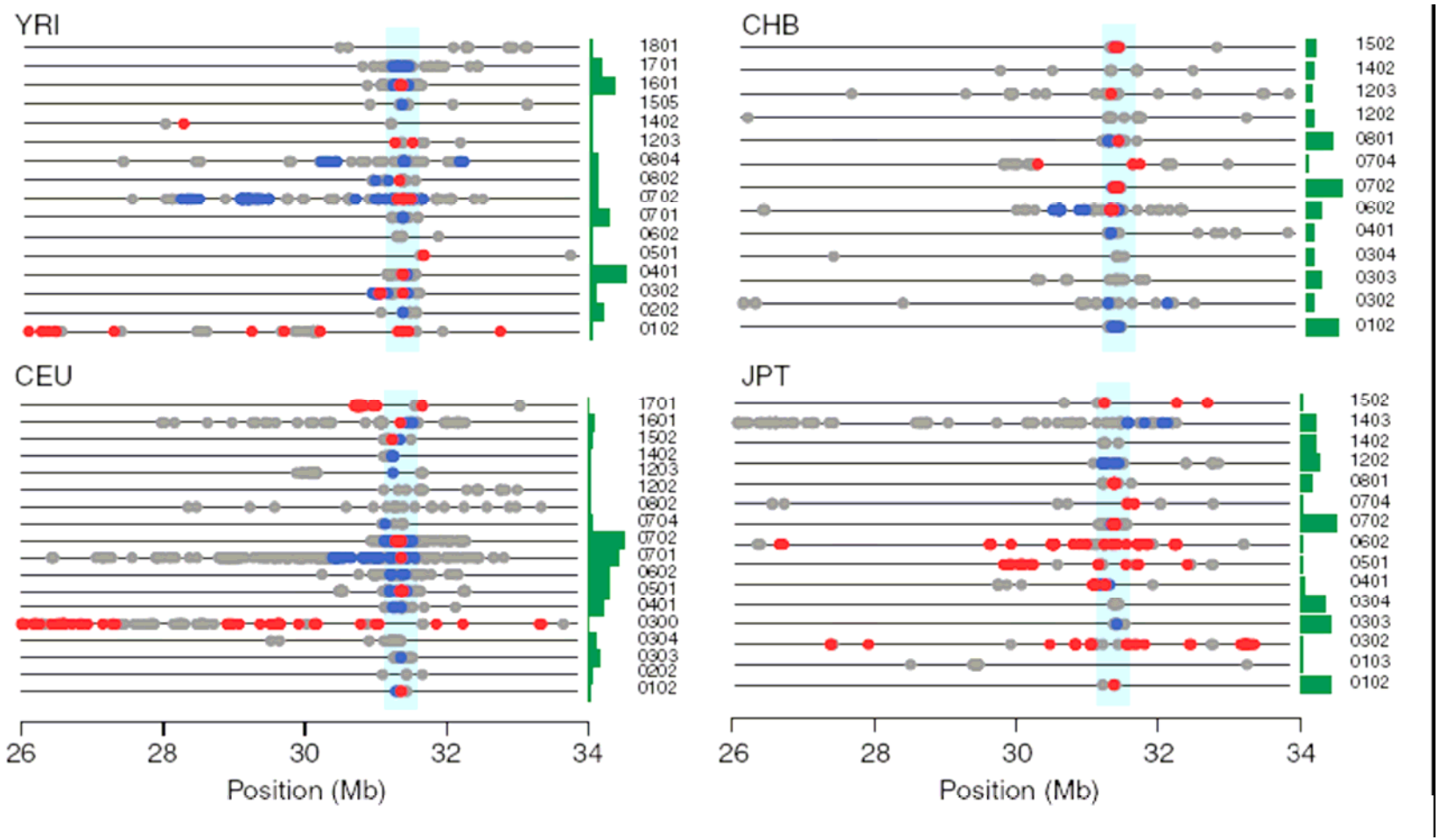
Haplotüüpide pikkused



The extent of association



Allelic association (HLA-C)



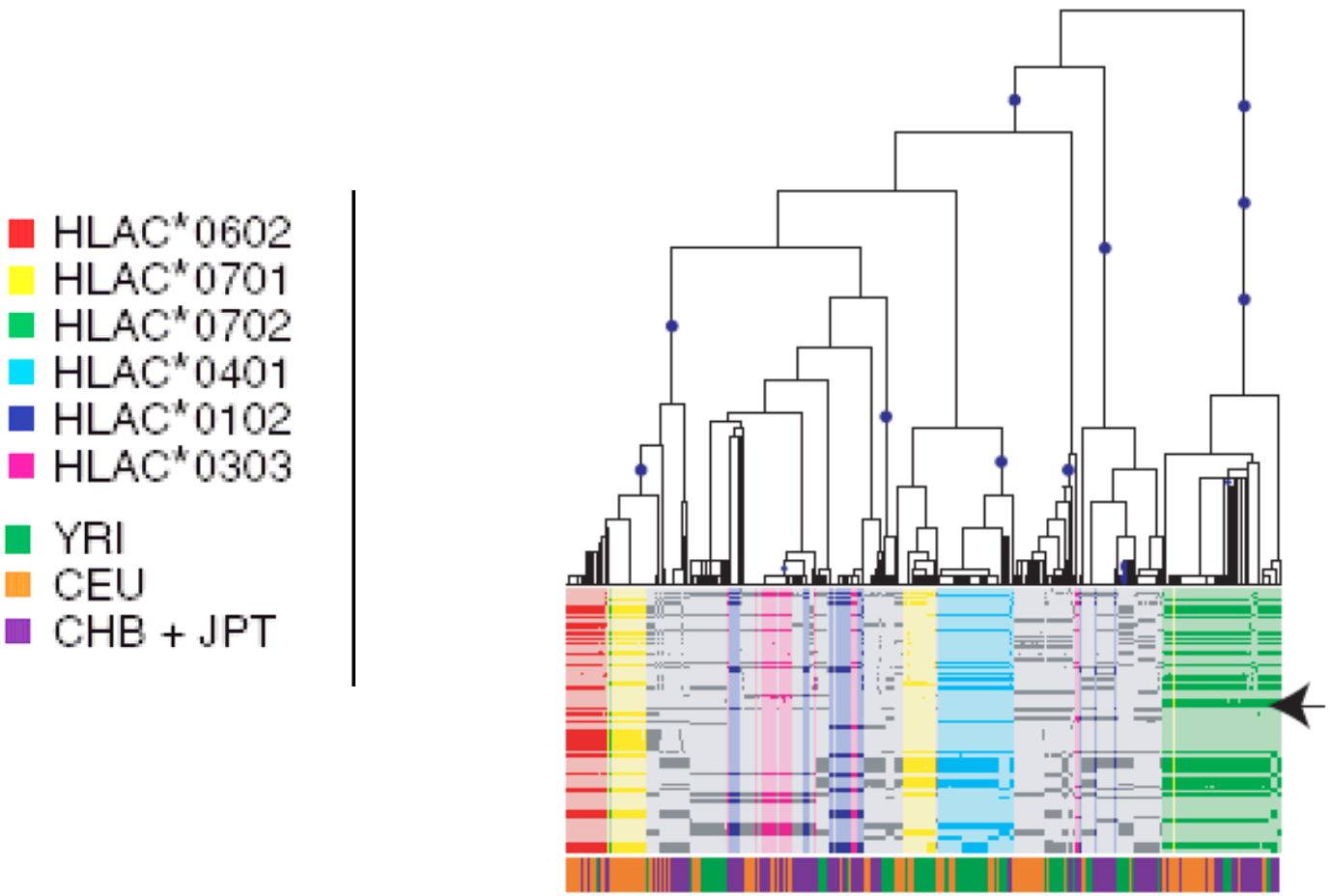
Empirical validation of SNP tags

Table 2 Empirical validation of SNP-based tags of associated HLA alleles in large collections of affected individuals

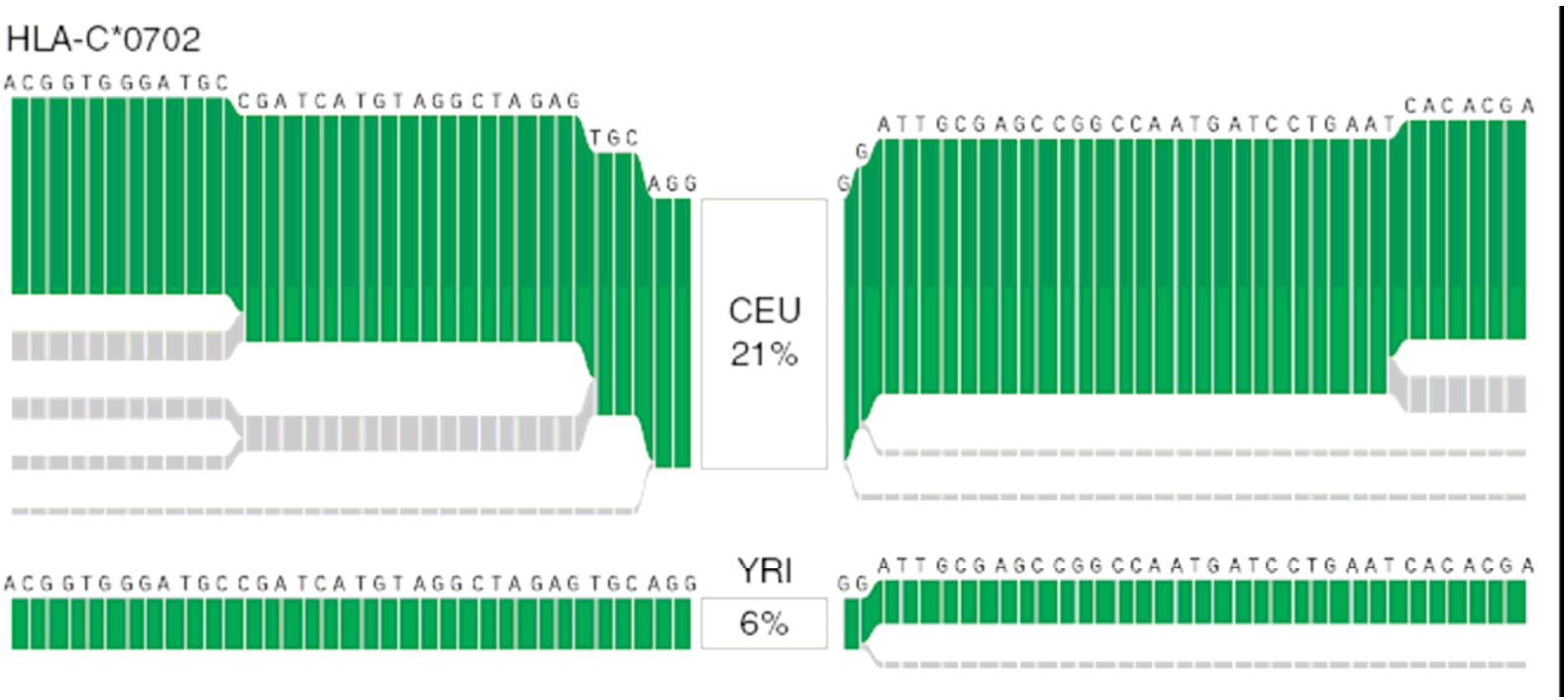
		Test based on SNP or haplotype					
	HLA allele	+	–	Total	Sensitivity (%)	Specificity (%)	r^2
Celiac disease							
DQA1*0201–DQB1*0202 (DQ2.2) / DQA1*0501–DQB1*0201 (DQ2.5) heterozygote	+	56	2	58			
	–	1	271	272			
	Total	57	273	330	96.6	99.6	0.94
DQA1*0501–DQB1*0201 (DQ2.5) homozygote	+	72	3	75			
	–	3	252	255			
	Total	75	255	330	96.0	98.8	0.90
SLE							
DRB1*1501	+	161	6	167			
	–	12	1149	1161			
	Total	173	1155	1328	96.4	99.0	0.88
DRB1*0301	+	245	5	250			
	–	24	1054	1078			
	Total	269	1059	1328	98.0	97.8	0.87

The test for DQ2.2 (DQA1*0201/DQB1*0202) is the rs4988889(T), rs2858331(C) haplotype; the test for DQ2.5 (DQA1*0501/DQB1*0201) is the rs4988889(T), rs2858331(T) haplotype; the test for DRB1*1501 is rs3135388; and the test for DRB1*0301 is rs2187688.

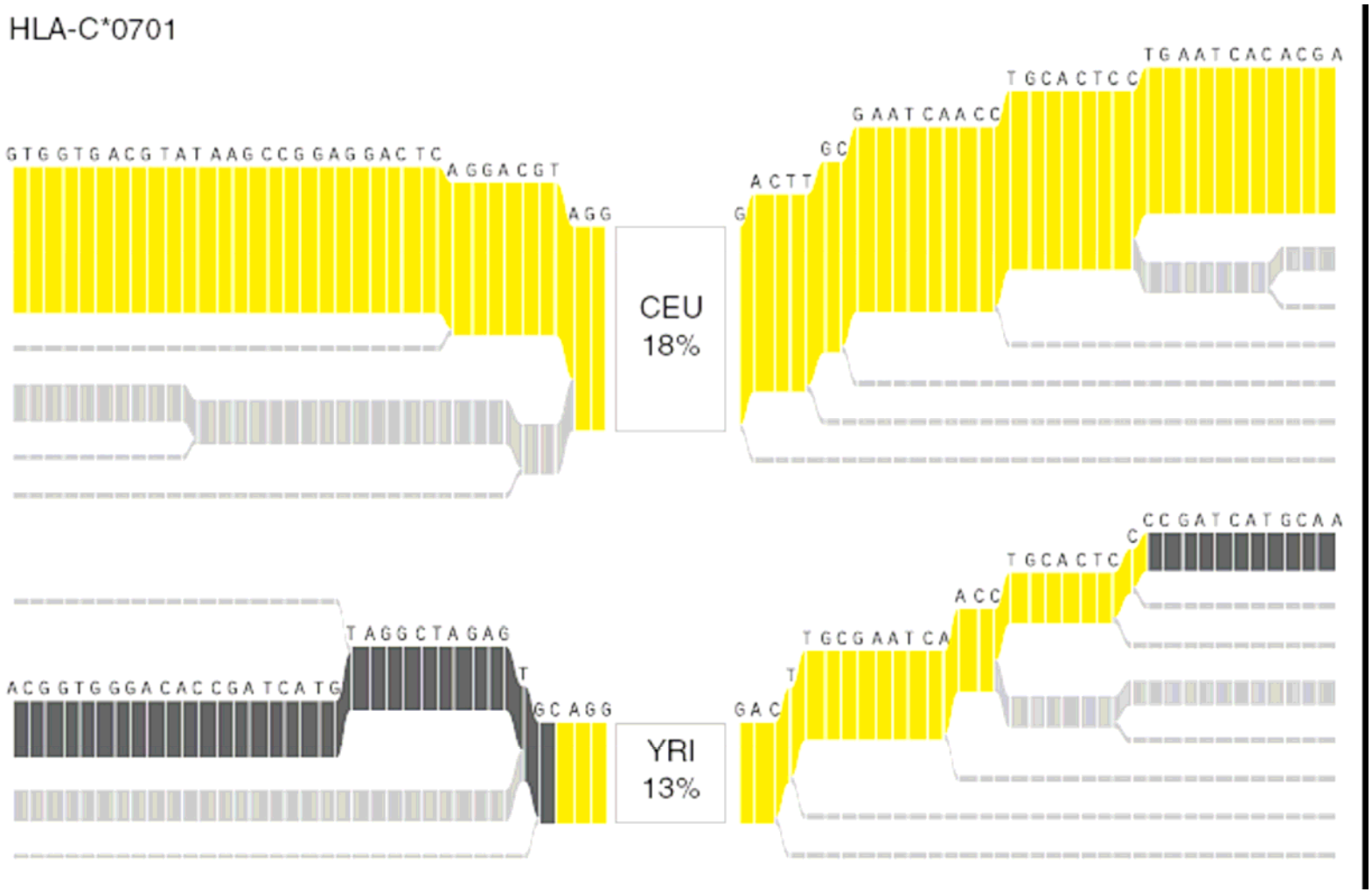
The evolutionary history of HLA-C



The evolutionary history of HLA-C



The evolutionary history of HLA-C



Long-range haplotype association