

INMUNOINFORMATICA

Pedro A Reche, PhD

INDEX

- **IMMUNE SYSTEM**
- **AREAS OF RESEARCH IN IMMUNOINFORMATICS**
- **COMPUTATIONAL VACCINOLOGY**

INMUNOINFORMATICA?

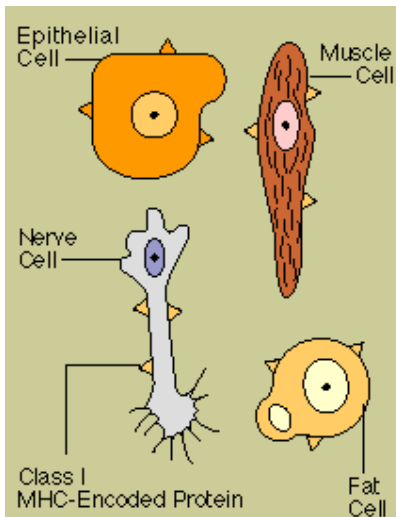
- IMMUNOINFORMATICS: Study of the immune system through the use of computers
 - Existencia masiva de Datos concernientes al sistema inmune
 - Problemas especificos relacionados con el sistema inmune
 - **Complejidad del sistema**

IMMUNE SYSTEM FUNCTION: SELF VS NON SELF RECOGNITION

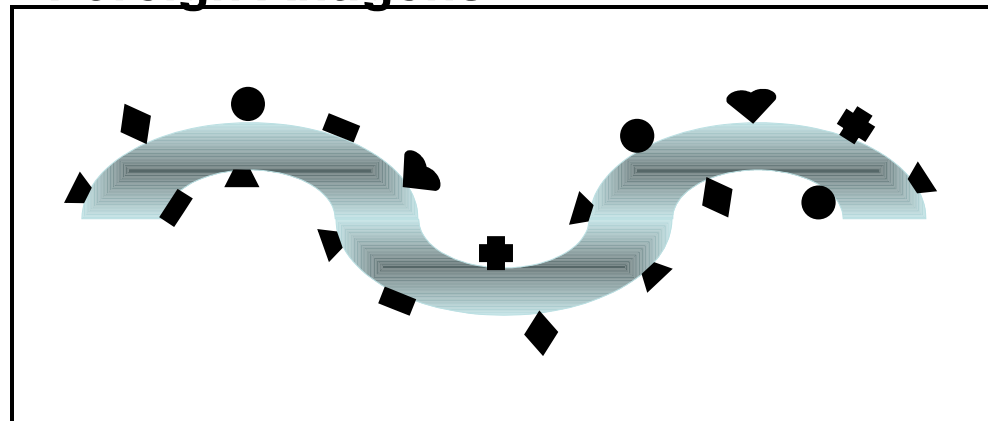
The immune system offer protection against infections

- Immune system function is based on discriminating between self and non-self
- Self and non-self discrimination is achieved through the recognition of small molecular subunits (antigens): self and non-self antigens
- Self antigens confer tolerance and non-self (foreign) antigens activate the immune system

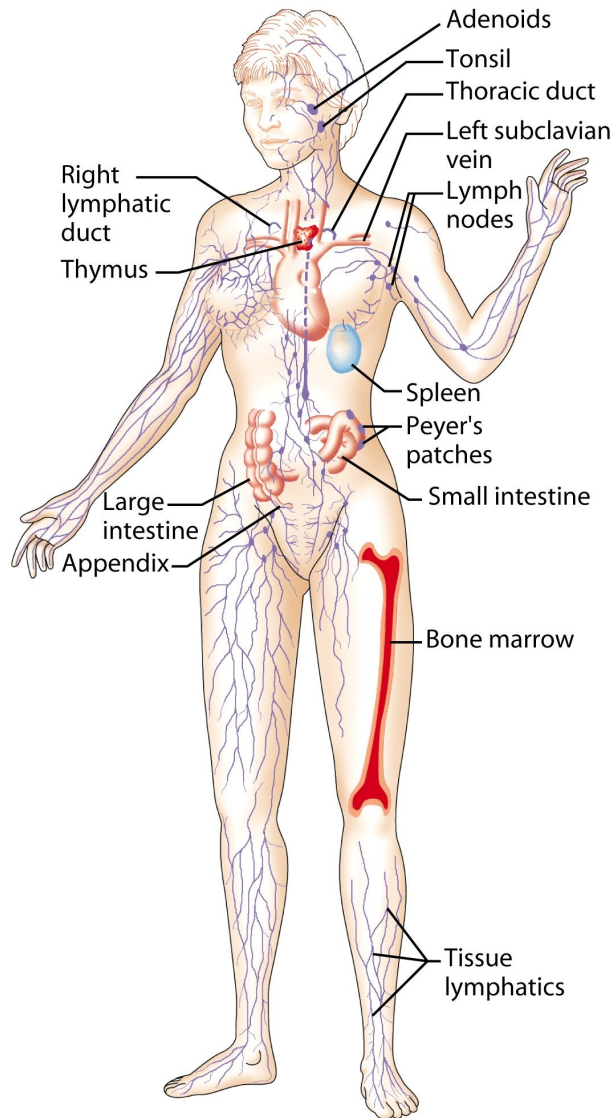
Self Antigens



Foreign Antigens



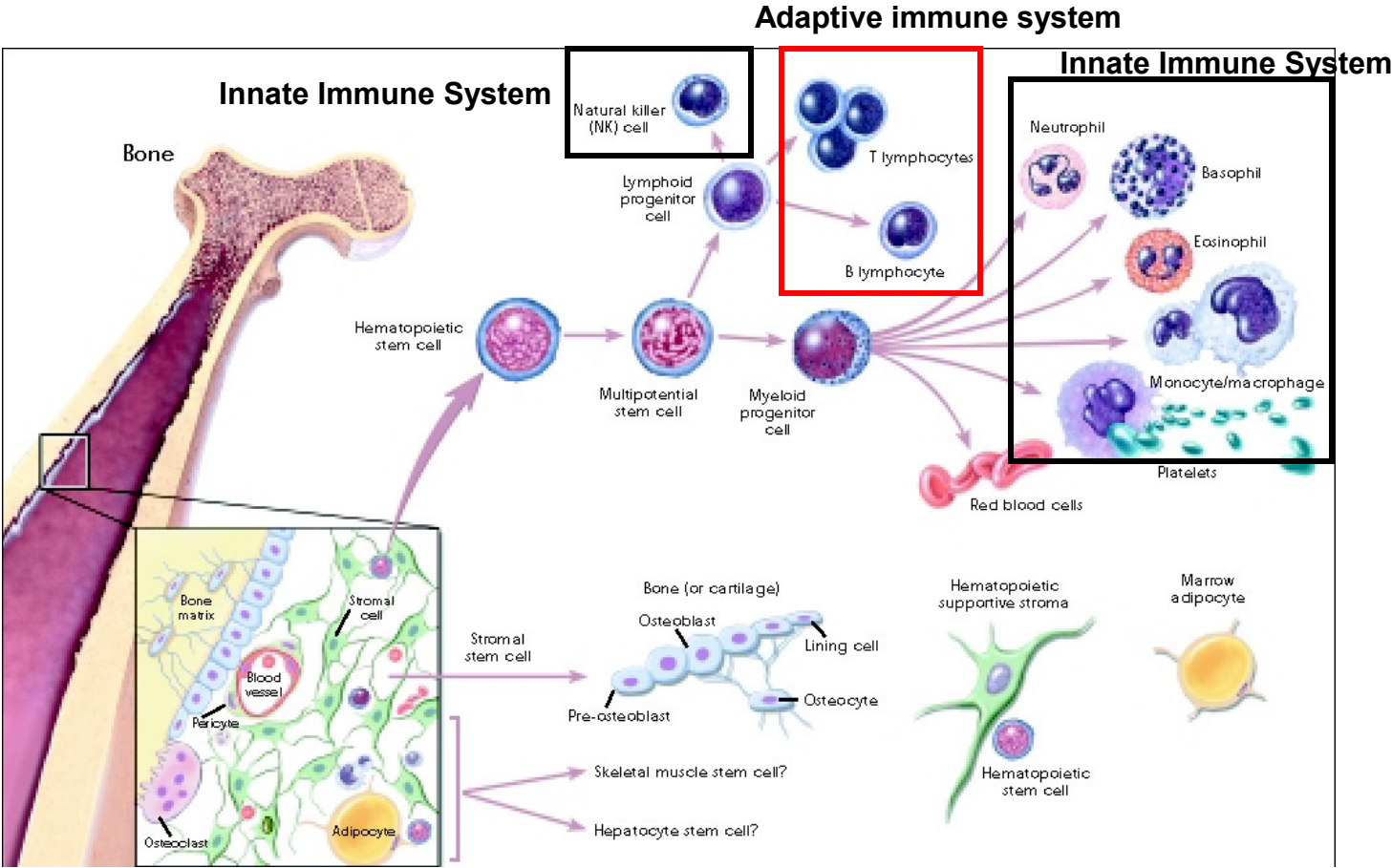
IMMUNE SYSTEM ORGANS



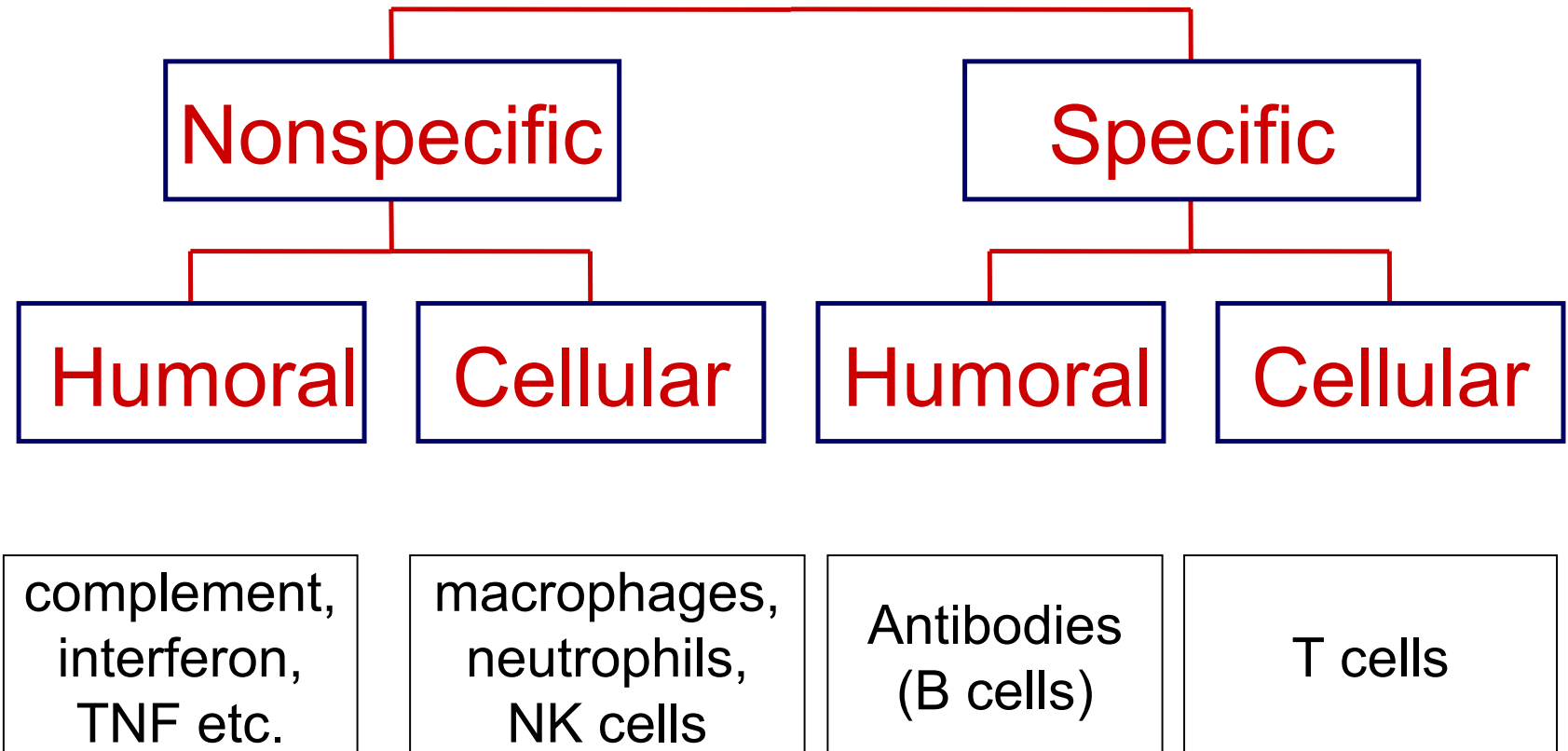
Functions:

- Production of Immune system cells
- Maturation and education of Immune system cells
- Immune system recognition and activation

IMMUNE SYSTEM CELLS: HEMATOPOYESIS

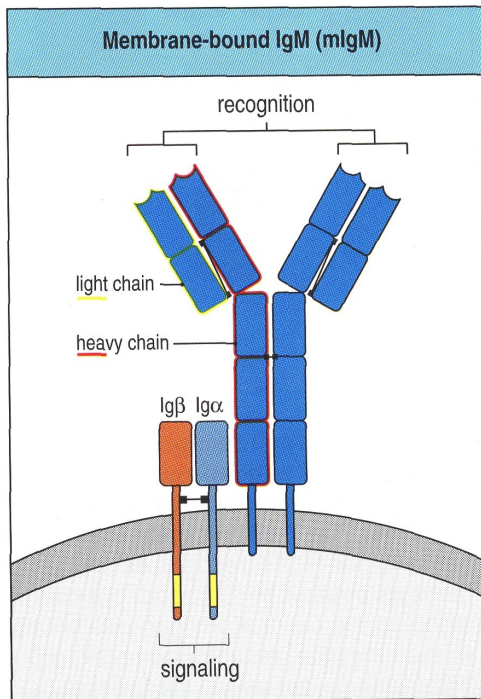


Innate versus Adaptive Immune System



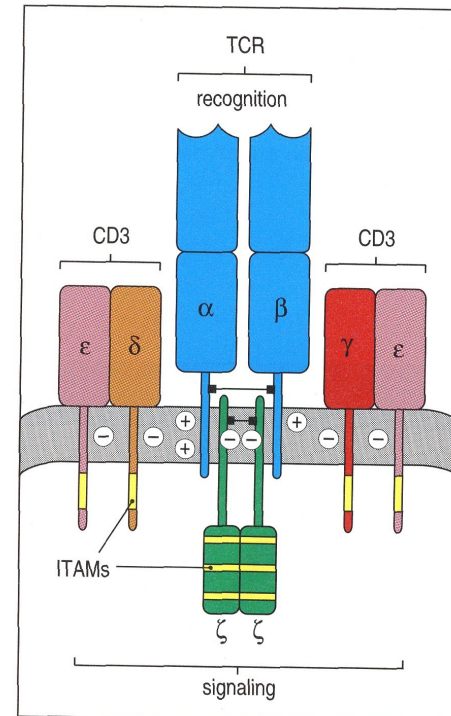
ANTIGEN SPECIFIC RECEPTORS

BCR



10^{20}

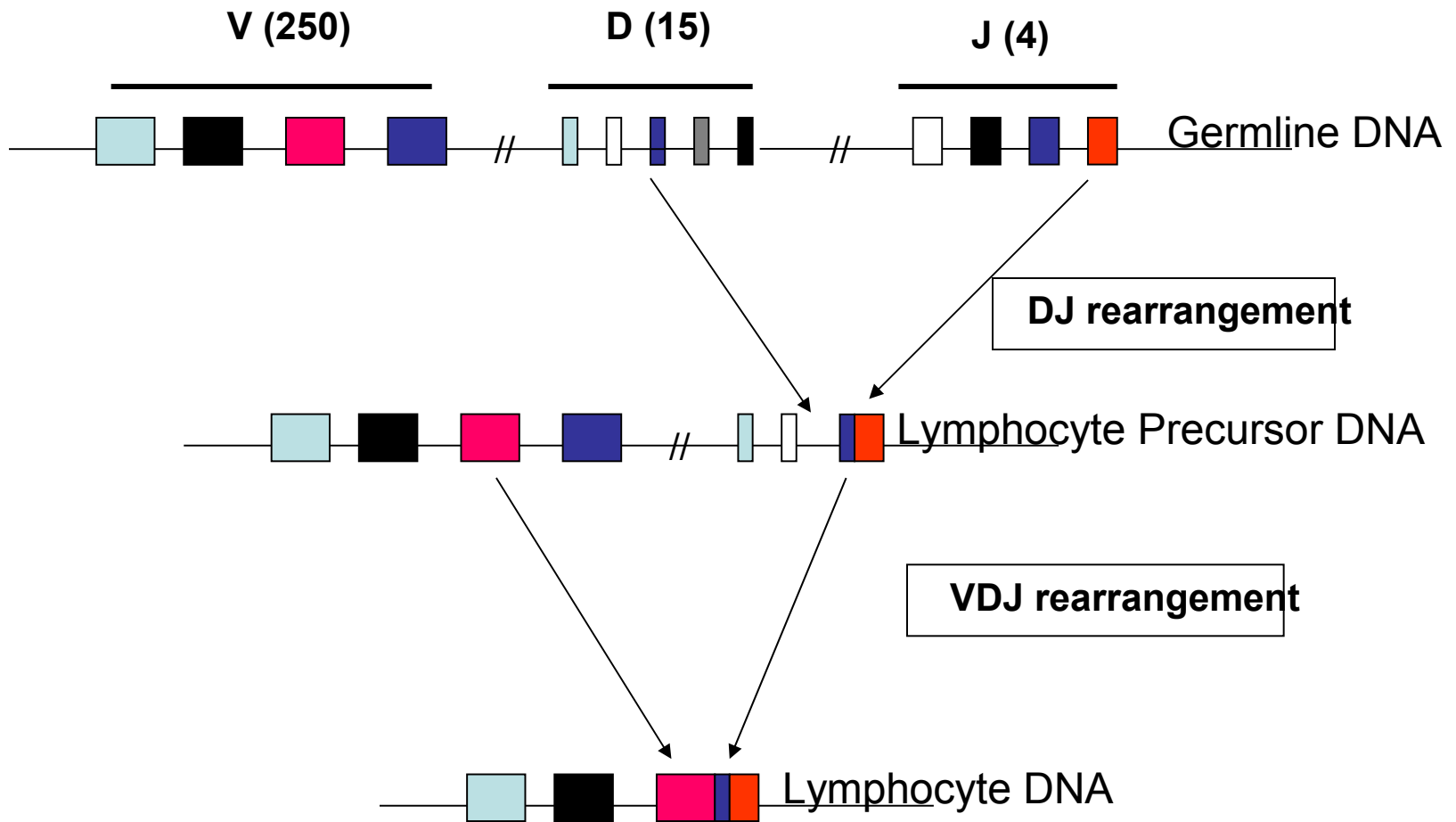
TCR



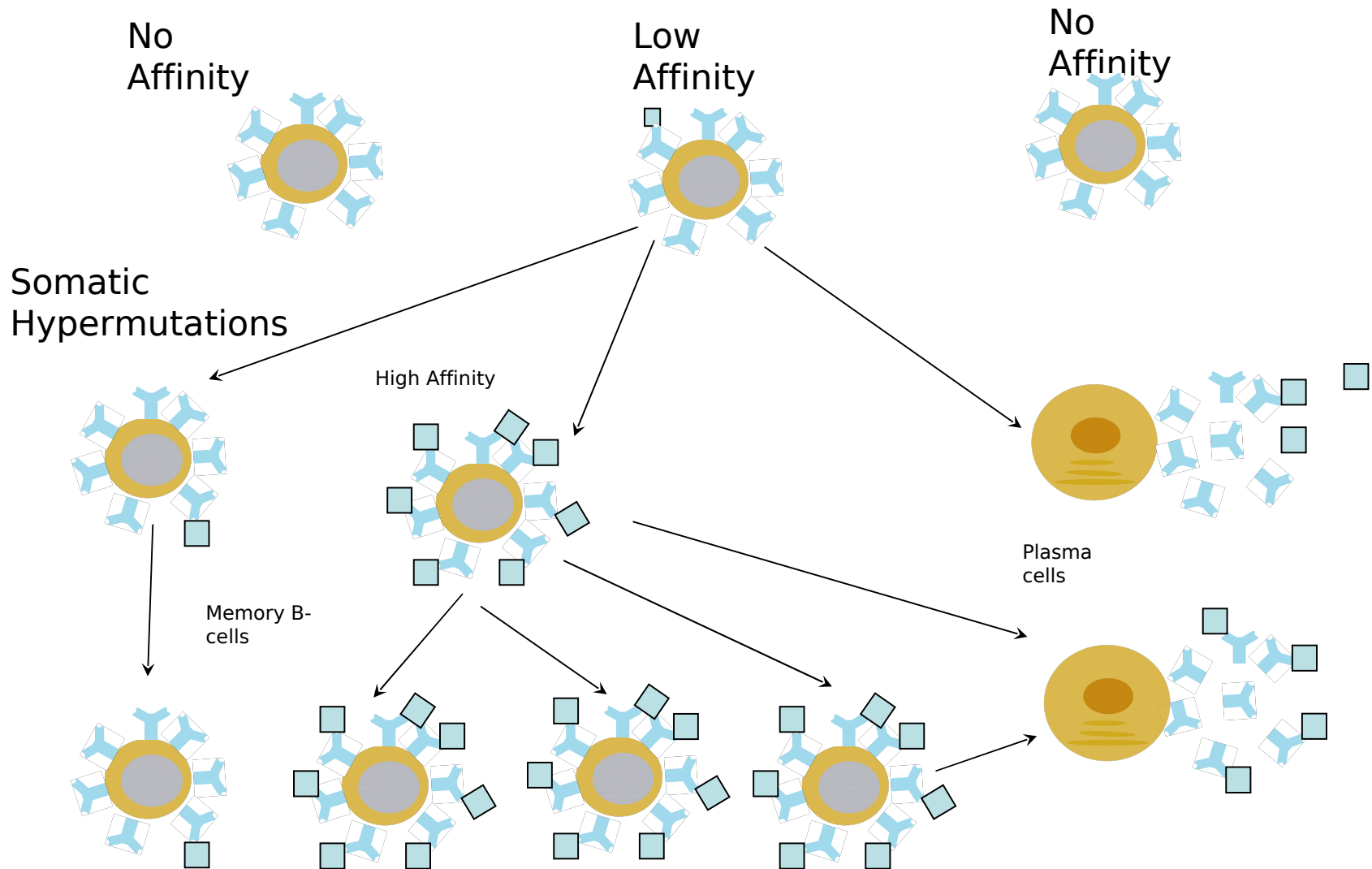
10^8

DIVERSITY OF TCR AND BCR REPERTOIRES IS GENERATED BY SOMATIC RECOMBINATION

(S. Tonegawa)

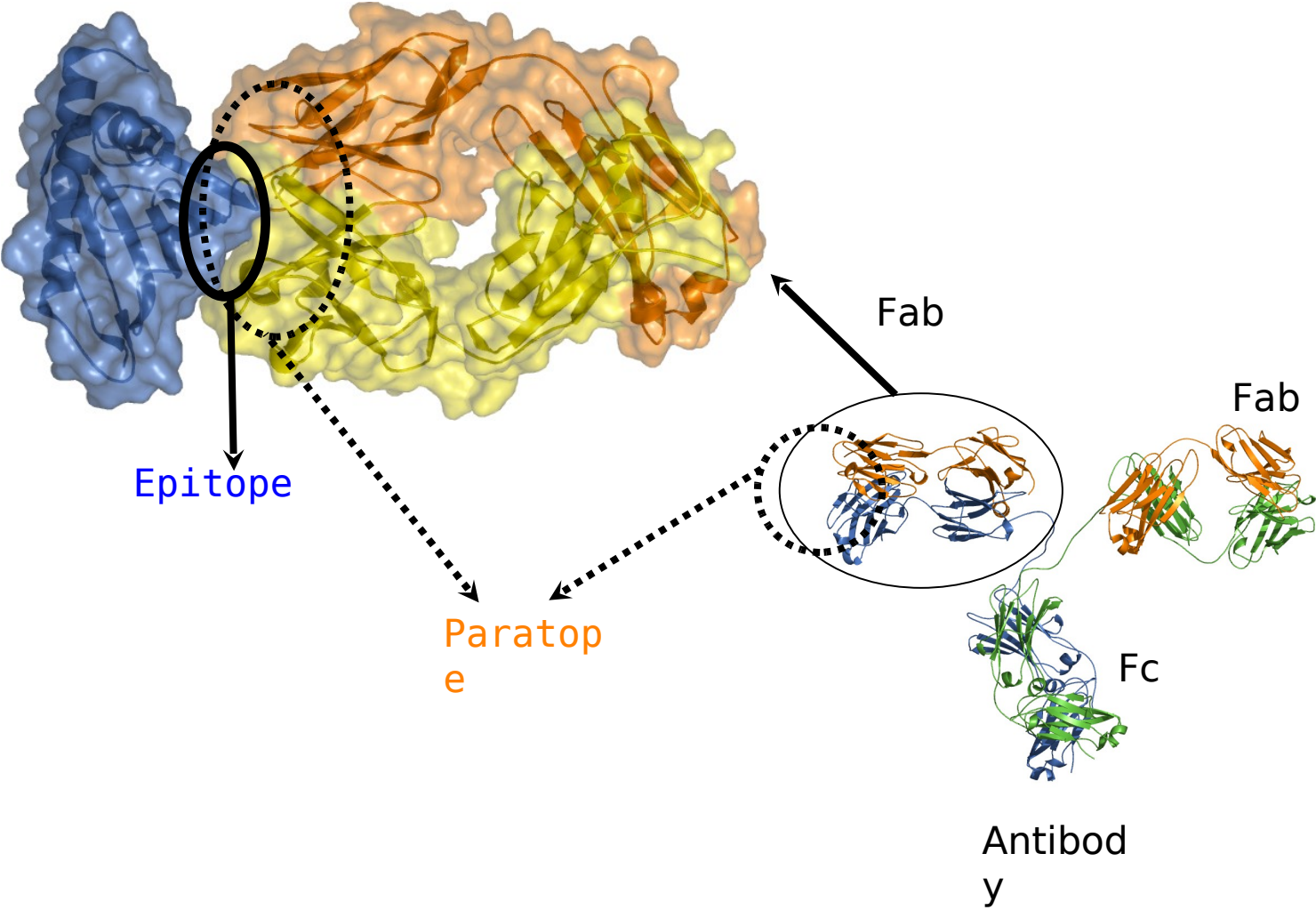


B-CELL ACTIVATION: SELECTION AND CLONAL EXPANSION



ANTIBODY - ANTIGEN INTERACTION

Antigen



ANTIGEN RECOGNITION BY T CELLS

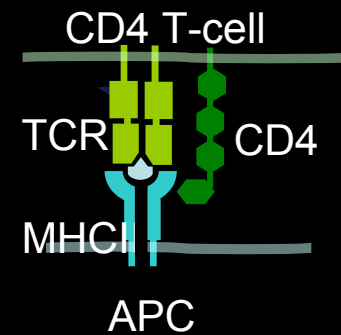
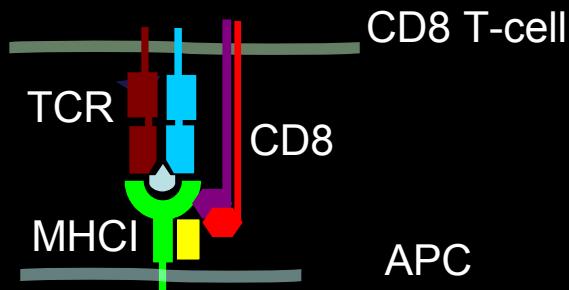
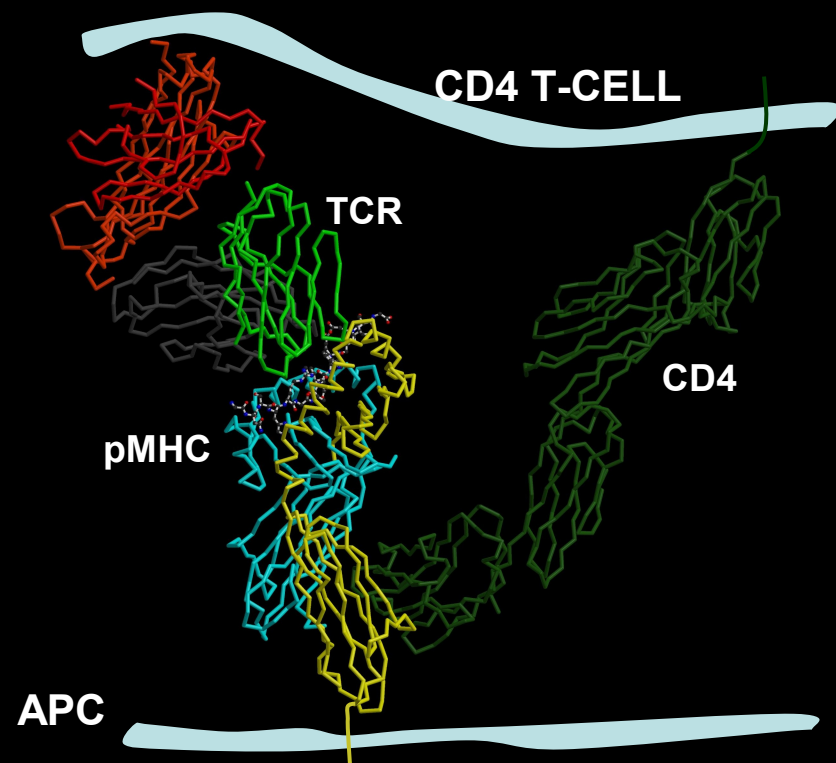
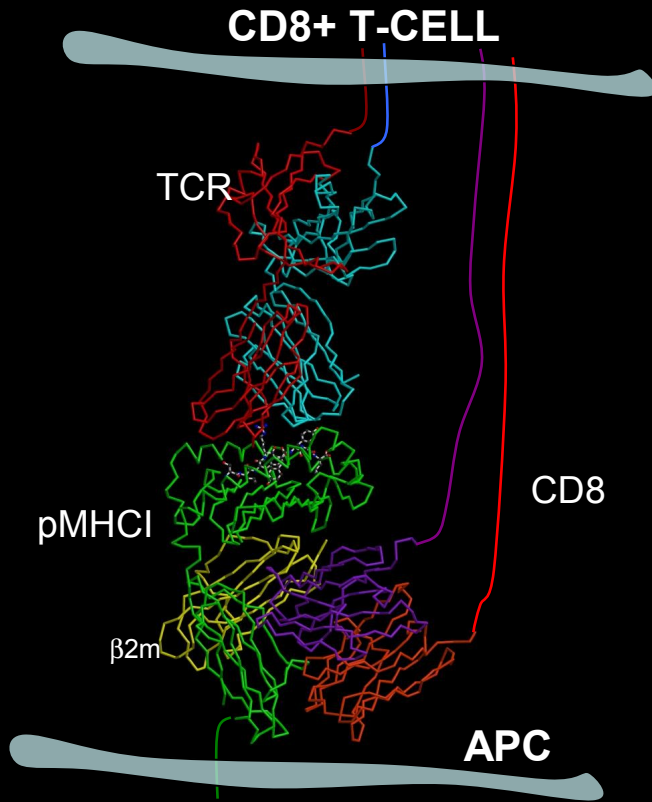
1 : T cells recognize peptides (about 9 amino acids) and not native antigen

2 : T cells recognize peptides when bound to their own MHC

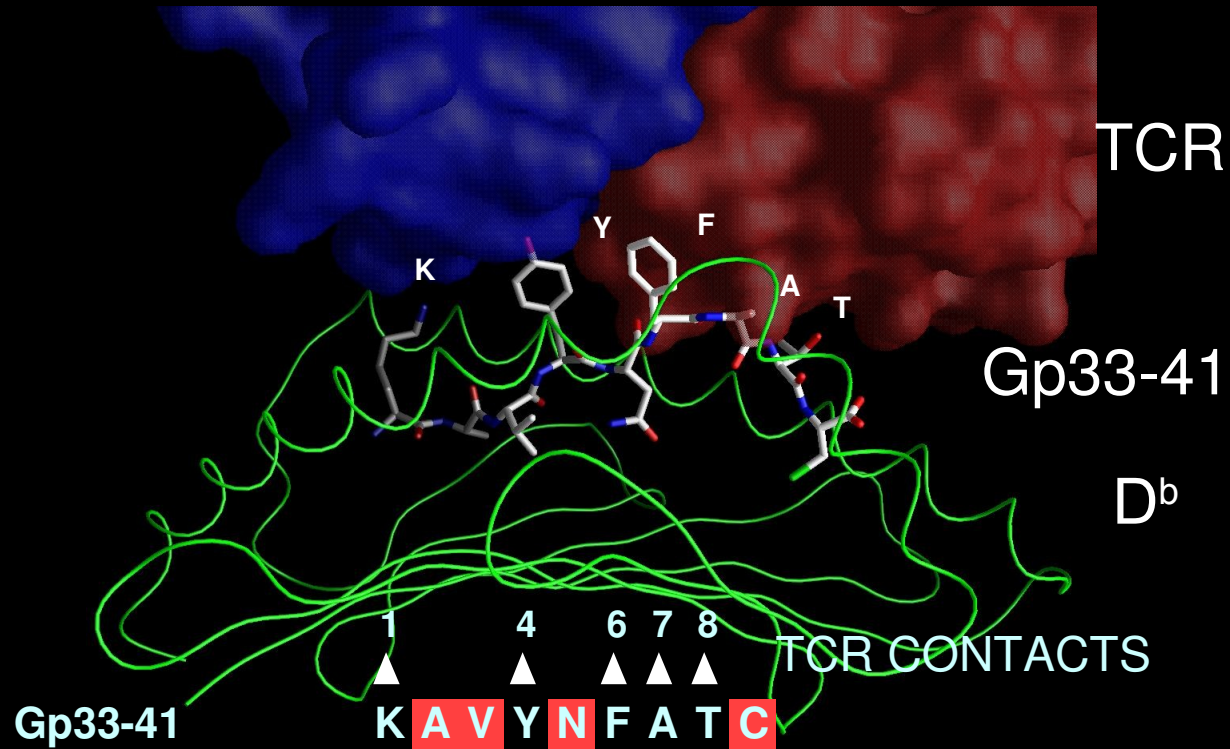
CD4 T cells : MHC-cl II

CD8 T cells : MHC-cl I

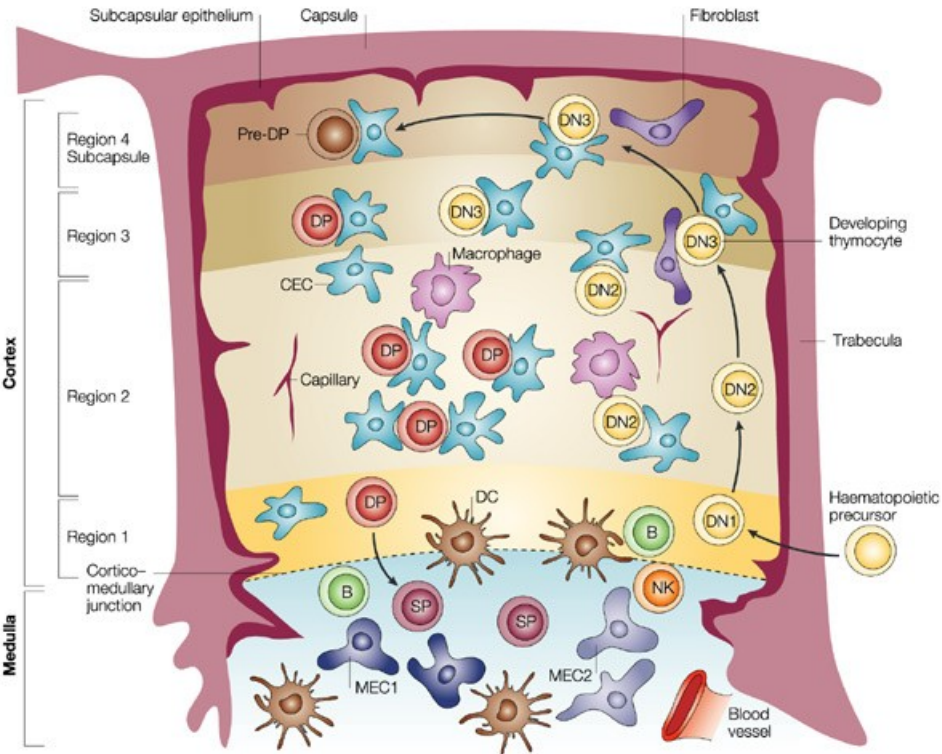
Structural recognition of peptide-MHCI by TCR



Structural features of TCR Recognition of pMHCII complexes



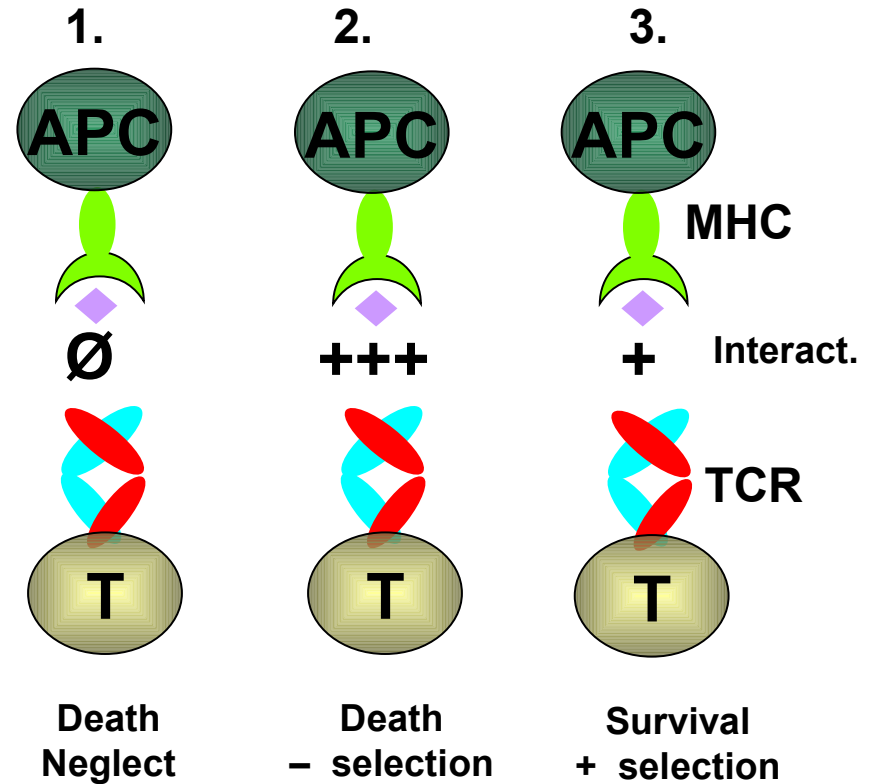
TCR-PMHC FIT AND SELF TOLERANCE OF T-CELLS IS ADQUIRED DURING THYMIC SELECTION



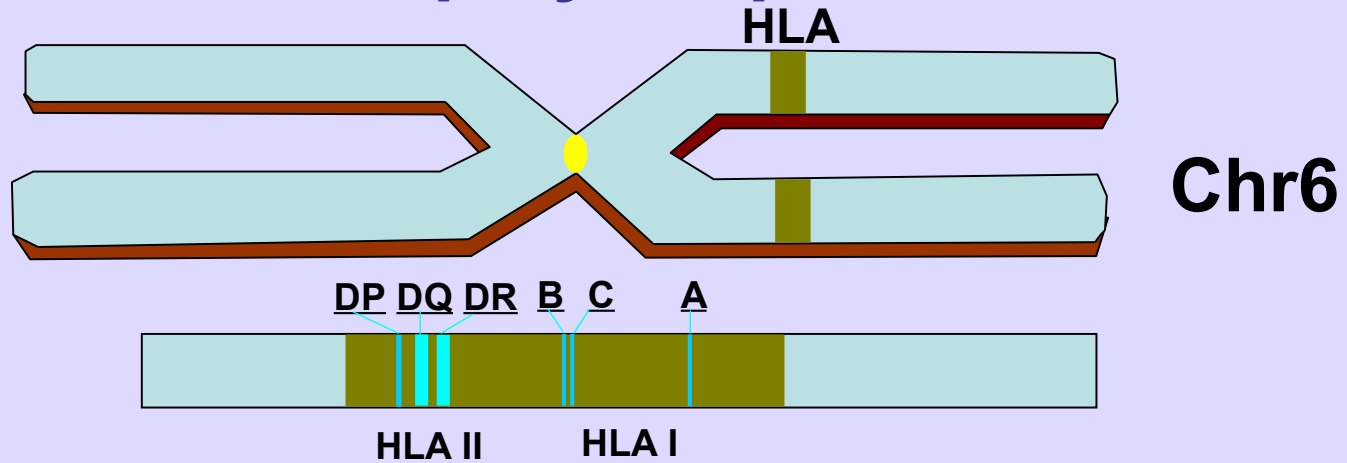
Nature Reviews | Immunology

Nature Reviews Immunology 4; 278-289 (2004);

Thymocyte selection is mediated by peptide/MHC/TCR interactions



In humans, MHC molecules are extremely polymorphic



HLA MOLECULE	SEQUENCES
CLASS I	
HLA-A	230
HLA-B	447
HLA-C	97
CLASS II	
HLA-DPA	12
HLA-DPB	90
HLA-DQA	17
HLA-DQB	42
HLA-DRA	2
HLA-DRB1	271
HLA-DRB3	30
HLA-DRB4	7
HLA-DRB5	14

	Black	Caucas .	Hispan .	Nat .Ame	Asian
GF	5.6%	15.1%	6.0%	7.5%	1.5%
PF	10.8%	27.9%	11.6%	14.4%	3.0%
HLA I					
	Black	Caucas .	Hispan .	Nat .Ame	Asian
GF	2.0%	11.75%	6.7%	3.4%	0.7%
PF	3.9 %	22.0 %	12.39%	8.3 %	1.3 %

INNATE AND ADAPTIVE IMMUNE SYSTEM

ANTIGEN PROCESSING AND PRESENTATION

INNATE



ADAPTIVE

CELLS OF THE INNATE IMMUNE SYSTEM WORK AS ANTIGEN PRESENTING CELLS

Phagocytic cells:

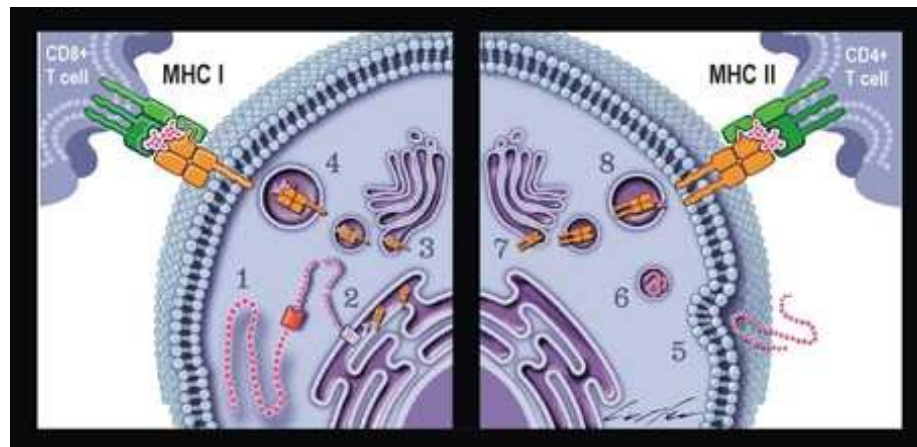
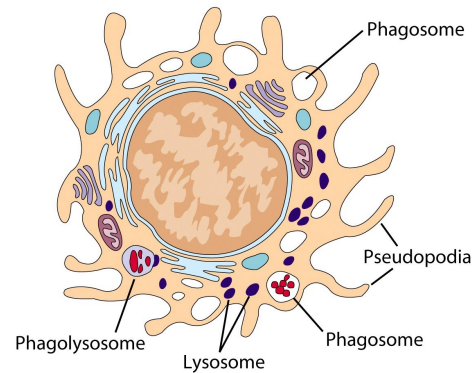
Granulocytes

Monocytes

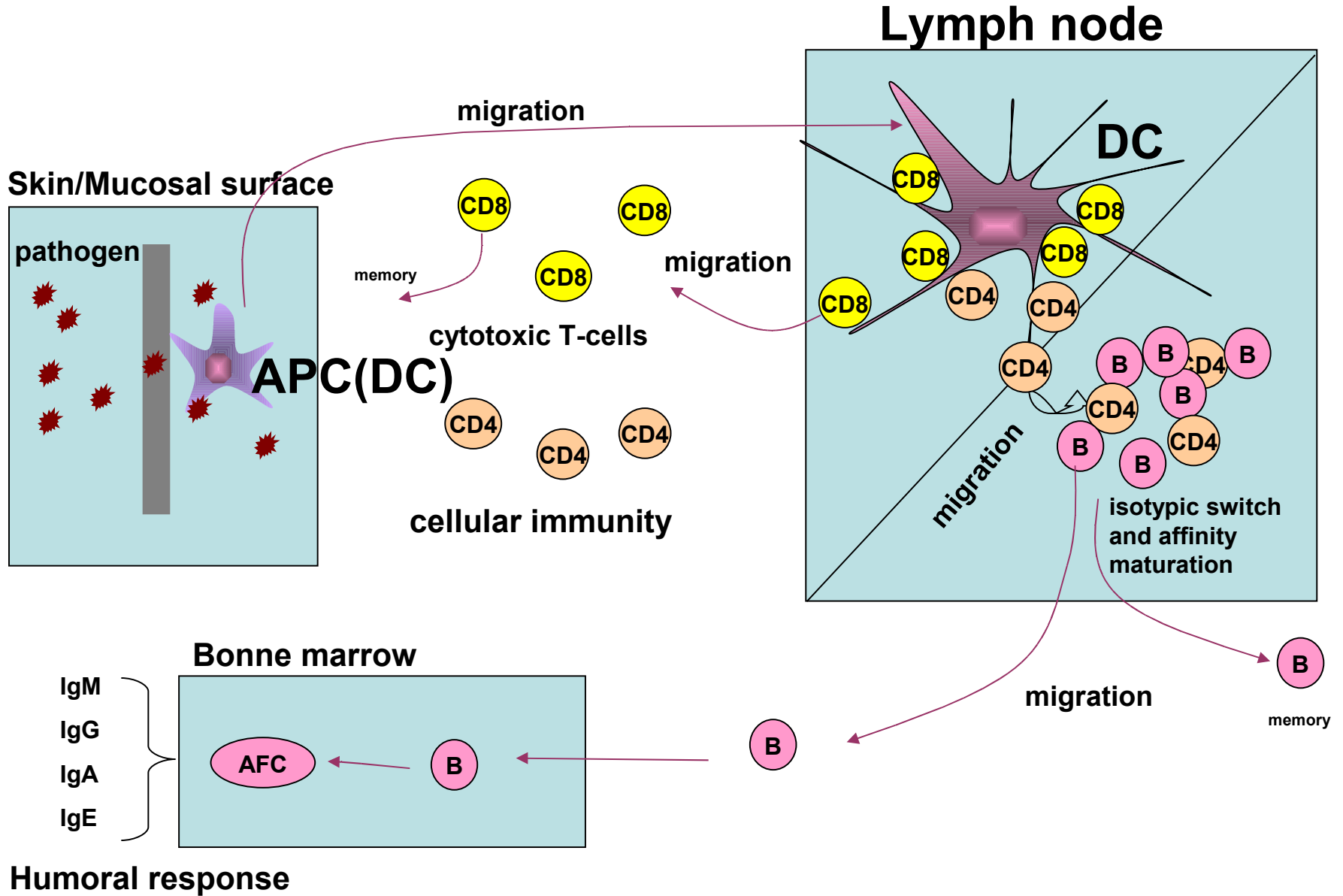
Macrophages

Dendritic cells

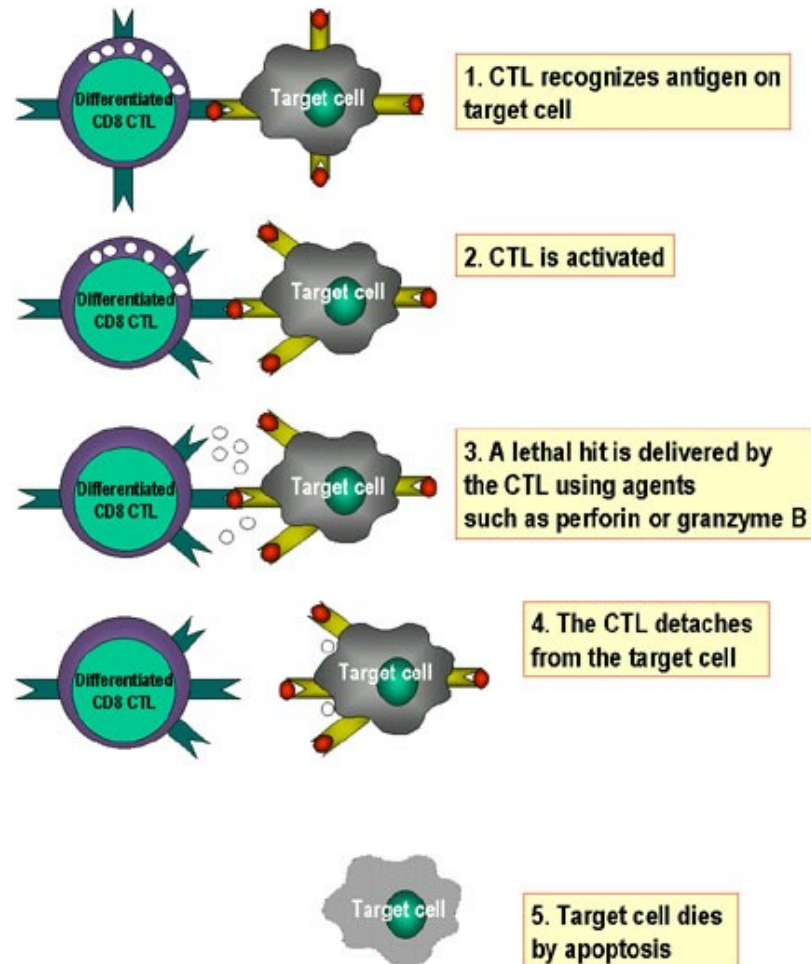
(b) Macrophage



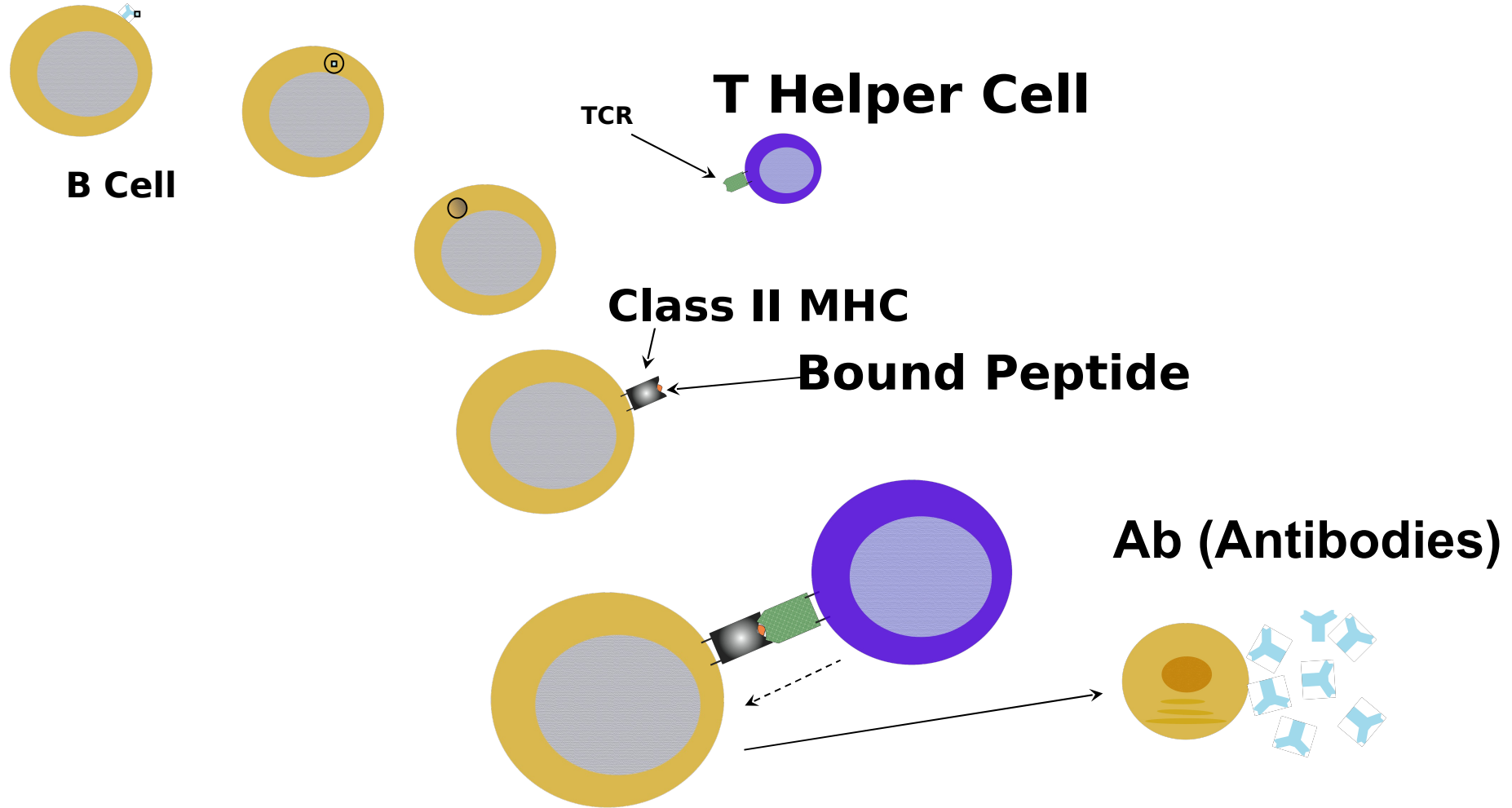
IMMUNE RESPONSE



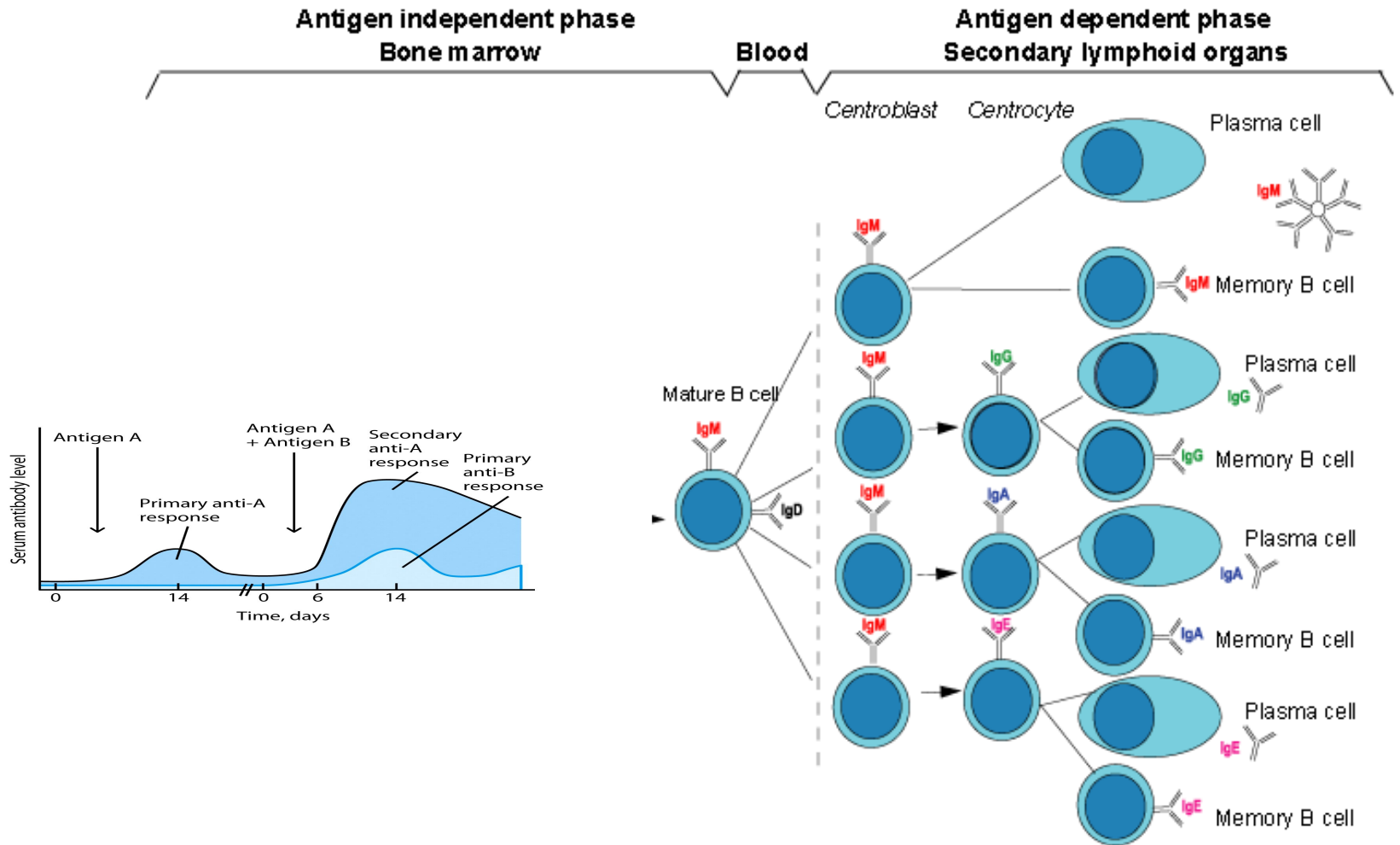
CD8 T cell Effector Function



CD4 T cell Effector Function: B-Cell Activation



B cells – memory cells



IMMUNOINFORMATICS: AREAS OF DEVELOPEMENT

- **Specialized Databases and Resources**
- **System level models describing immune system proceses**
- **Molecular models describing immune system features: Antigen recognition**

IMMUNOINFORMATICS: AREAS OF DEVELOPEMENT

- Generation of Specialized Databases and Resources
 - Databases of immunological sequences: IMGT
 - Databases of epitopes: EPIMHC
 - Database of pathogens: HIV database

IMMUNOINFORMATICS DATABASES: IMGT

<http://imgt.cines.fr/>

WELCOME !
to the IMGT Home page

THE INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



IMGT®, the international ImMunoGeneTics information system® <http://imgt.cines.fr>, is a high-quality integrated knowledge resource specialized in the immunoglobulins (IG), T cell receptors (TR), major histocompatibility complex (MHC), immunoglobulin superfamily (IgSF), major histocompatibility complex superfamily (MhcSF) and related proteins of the immune system (RPI) of human and other vertebrate species, created in 1989 by Marie-Paule Lefranc ([Université Montpellier II, CNRS](http://www.univ-montpellier.fr)). IMGT, a European project since 1992, works in close collaboration with [EBI](http://www.ebi.ac.uk). IMGT consists of [sequence](#) databases (IMGT/LIGM-DB, a comprehensive database of IG and TR from human and other vertebrates, with translation for fully annotated sequences, IMGT/MHC-DB, IMGT/PRIMER-DB), [genome](#) database (IMGT/GENE-DB) and [structure](#) database (IMGT/3Dstructure-DB). **Web resources** (IMGT Marie-Paule page) and **interactive tools**. The IMGT Home page <http://imgt.cines.fr> (Montpellier, France) provides a common access to all Immunogenetics data.

IMGT founder and director: [Marie-Paule Lefranc \(Marie-Paule.Lefranc@igh.cnrs.fr\)](mailto:Marie-Paule.Lefranc@igh.cnrs.fr), Université Montpellier II, CNRS, [LIGM](#), [IGH](#), [IFR3](#), Montpellier (France)

[IMGT® Site Map](#)
Information on IMGT® (creations and updates, references, FAQ, citing IMGT, funding support...)



IMGT databases

- [IMGT/LIGM-DB](#) (IG and TR from 204 species) (LIGM, Montpellier, France) (**107 367 entries**)
- [IMGT/MHC-DB](#) (IMGT/MHC-HLA, -NHP, -DLA, -FLA) (ANRI, BPRC, hosted at EBI)
- [IMGT/PRIMER-DB](#) (IG and TR from 11 species) (LIGM, Montpellier, France) (**1 864 entries**)
- [IMGT/GENE-DB](#) (IG and TR genes from human and mouse) (LIGM, Montpellier, France) (**1 511 genes, 2 462 alleles**)
- [IMGT/3Dstructure-DB](#) (IG and TR, MHC and RPI gene and allele identification and IMGT Colliers de Perles) (**1 269 entries**)

IMGT tools

- [IMGT/V-QUEST](#) (sequence alignment software for IG, TR and HLA)
- [IMGT/JunctionAnalysis](#) (for human and mouse IG and TR)
- [IMGT/Allele-Align](#)
- [IMGT/PhyloGene](#)
- [IMGT/DomainDisplay](#) (Amino acid sequences)
- [IMGT/LocusView](#), [IMGT/GeneView](#), [IMGT/GeneSearch](#), [IMGT/CloneSearch](#) (for human IGH, IGL, IGH, TRA/TRD, TRB, TRG, mouse TRA/TRD and human MHC)
- [IMGT/GeneInfo](#) (TIMC and ICH, Grenoble; LIGM, Montpellier)
- [IMGT/GeneFrequency](#)
- [IMGT/DomainGapAlign](#)
- [IMGT/Collier-de-Perles](#)
- [IMGT/DomainSuperimpose](#)
- [IMGT/StructuralQuery](#)

IMGT Web resources

- [IMGT Repertoire](#) (IG and TR, MHC and RPI)
- [IMGT Index](#) (FactsBook)
- [IMGT Scientific chart](#) (Sequence description, Numbering, Nomenclature, Representation rules)
- [IMGT Bloc-notes](#) (Interesting links, PubMed, Meeting announcements, Postdoctoral positions and jobs **NEW!**, Messages, Search engines...)
- [IMGT Education](#) (IMGT Lexique, Aide-mémoire, Tutorials, Questions and answers, Enseignements...)
- [IMGT Medical page](#), [IMGT Veterinary page](#), [IMGT Biotechnology page](#)
- [IMGT Posters and diaporama](#)
- [The IMGT Immunoinformatics page](#)

IMGT other accesses

- [IMGT Other accesses](#) (SRS, FTP...)
- [Compare your sequence against IMGT](#) (BLAST, FASTA)
- [IMGT/LIGM-DB Sequence submission](#)
- [IMGT flat file release information](#)

Search



WWW IMGT domain

IMMUNOINFORMATICS DATABASES: EPIMHC

<http://bio.dfci.harvard.edu/epimhc/>

Databases >> EPIMHC Database



EPIMHC

A Curated Database of MHC Ligands

<input type="checkbox"/> All <input type="checkbox"/> BOLA-A11 <input type="checkbox"/> DR52 <input type="checkbox"/> ELA-A5 <input type="checkbox"/> ELA-A9 <input type="checkbox"/> H2-AB <input type="checkbox"/> H2-AD <input type="checkbox"/> H2-AG7 <input type="checkbox"/> H2-AK <input type="checkbox"/> H2-AS <input type="checkbox"/> H2-AU <input type="checkbox"/> H2-DB <input type="checkbox"/> H2-DD <input type="checkbox"/> H2-DK <input type="checkbox"/> H2-DQ <input type="checkbox"/> H2-EB <input type="checkbox"/> H2-ED <input type="checkbox"/> H2-EC7 <input type="checkbox"/> H2-EK <input type="checkbox"/> H2-ES <input type="checkbox"/> H2-K8 <input type="checkbox"/> H2-KD <input type="checkbox"/> H2-KK <input type="checkbox"/> H2-LD <input type="checkbox"/> H2-M3 <input type="checkbox"/> H2-M3WT <input type="checkbox"/> H2-QA-1A <input type="checkbox"/> H2-QA-1B <input type="checkbox"/> H2-QA-2 <input type="checkbox"/> H2-QA-2A <input type="checkbox"/> H2-A <input type="checkbox"/> H2-B	<input type="button" value="AND"/> <input type="button" value="↓"/>	SEQ <input type="text"/>
	<input type="button" value="AND"/> <input type="button" value="↓"/>	LENGTH <input type="text" value="6"/> <input type="text" value="7"/> <input type="text" value="8"/>
	<input type="button" value="AND"/> <input type="button" value="↓"/>	CLASS <input type="text" value="All"/>
	<input type="button" value="AND"/> <input type="button" value="↓"/>	MHC SOURCE <input type="text" value="All"/> <input type="text" value="BONOBO CHIMPANZEE"/> <input type="text" value="CHIMPANZEE"/> <input type="text" value="COTTON-TOP TAMARIN"/>
	<input type="button" value="AND"/> <input type="button" value="↓"/>	PEPTIDE SOURCE ORGANISM <input type="text" value="All"/> <input type="text" value="Chlamydia trachomatis"/> <input type="text" value="Clostridium botulinum"/> <input type="text" value="Dengue virus"/>
	<input type="button" value="AND"/> <input type="button" value="↓"/>	PEPTIDE BINDING LEVEL <input type="text" value="All"/> <input type="text" value="HIGH"/> <input type="text" value="LITTLE"/> <input type="text" value="MODERATE"/>
	<input type="button" value="AND"/> <input type="button" value="↓"/>	IMMUNOGENIC PEPTIDES (EPITOPES) <input type="text" value="All"/>
	<input type="button" value="AND"/> <input type="button" value="↓"/>	IMMUNOGENICITY LEVEL <input type="text" value="All"/> <input type="text" value="HIGH"/> <input type="text" value="LITTLE"/> <input type="text" value="MODERATE"/>
	<input type="button" value="AND"/> <input type="button" value="↓"/>	PROCESSING: <input type="text" value="All"/>
	RESULTS DISPLAY	
Select Fields to be Displayed	<input type="text" value="DEFAULT"/> <input type="text" value="MHC MOLECULE"/> <input type="text" value="SEQUENCE"/> <input type="text" value="MHC Source"/> <input type="text" value="CLASS"/> <input type="text" value="SEQUENCE LENGTH"/>	
Select field to order results by:	<input type="text" value="MHC"/>	<input type="text" value="Ascending"/>
<input type="button" value="Reset"/>	<input type="button" value="Search"/>	

IMMUNOINFORMATICS: SYSTEM LEVEL MODELS

- **Symbolic models** describing processes at cellular and organ levels:
 - **Immune response to viruses.** Dahari H, Lo A, Ribeiro RM, Perelson AS. Modeling hepatitis C virus dynamics: liver regeneration and critical drug efficacy. *J Theor Biol.* 2007 Jul 21;247(2):371-81.
 - **Analysis of MHC diversity under host-pathogen coevolution.** Bhattacharya T, et al effects in the assessment of HIV polymorphisms and HLA allele associations. *Science.* 2007 Mar 16;315(5818):1583-6.
 - **B cell maturation.** Meyer-Hermann et al. An analysis of B cell selection mechanisms in germinal centers. *Math Med Biol.* 2006 Sep;23(3):255-77
 - **Dynamic models that simulate both cellular and humoral immune responses.** Pappalardo et al simulation of cancer immunoprevention vaccine. *Bioinformatics.* 2005 Jun 15;21(12):2891-7.

IMMUNOINFORMATICS: MOLECULAR MODELS

- **Use molecular data (sequences, 3D etc)**
 - B cell epitope prediction
 - Modeling of antigen processing events
 - **T cell epitope prediction**
 - **Computer aided design of vaccines: Formulation of vaccines**

WHY BOTHER EPITOPE PREDICTION?

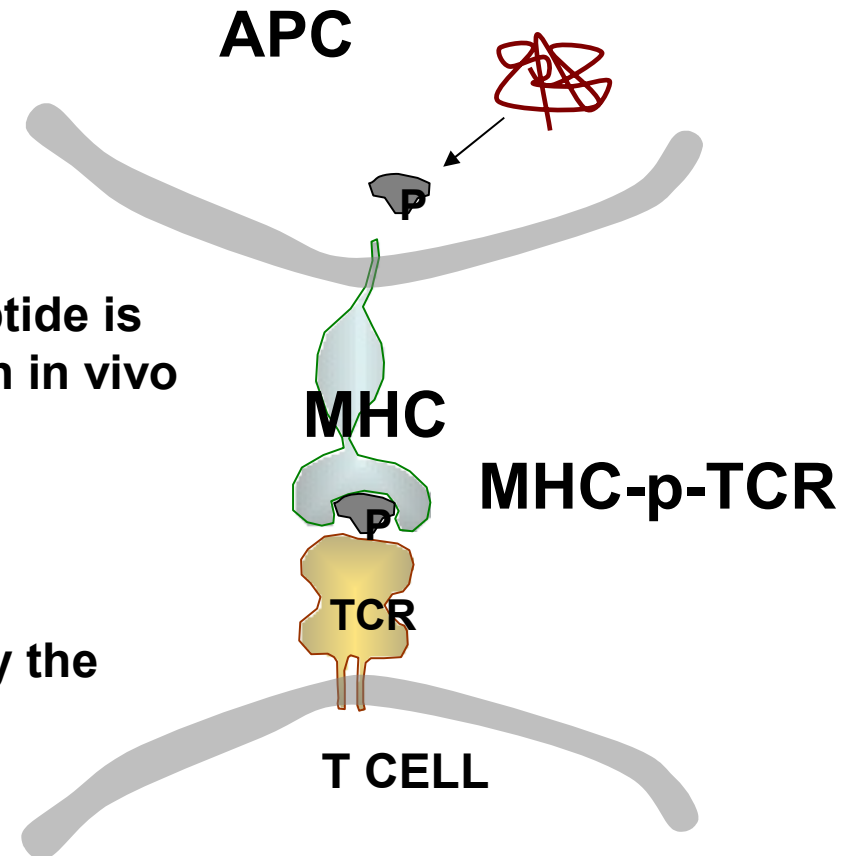
- Predicting the target of the immune system is of practical relevance for:
 - Understanding disease pathogenesis and for making vaccines
 - Understanding the processes selection tolerance, memory and aging of the immune system

Prediction of T cell epitopes

⇒ T cell epitopes are peptides derived from the processing of proteins that are able to activate T cells in a detectable manner

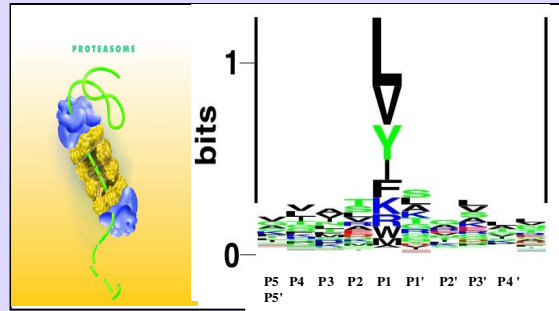
Mechanisms that needs to be modeled

- Antigen processing. Whether a peptide is processed (released) from a protein in vivo
- Binding of the peptide to the MHC molecule
- Recognition of the peptide-MHC by the TCR with enough affinity.

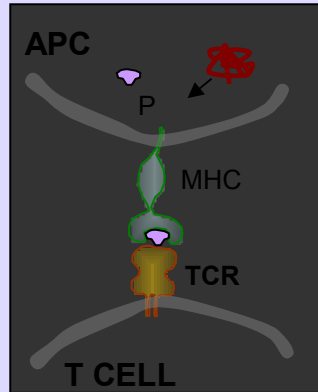
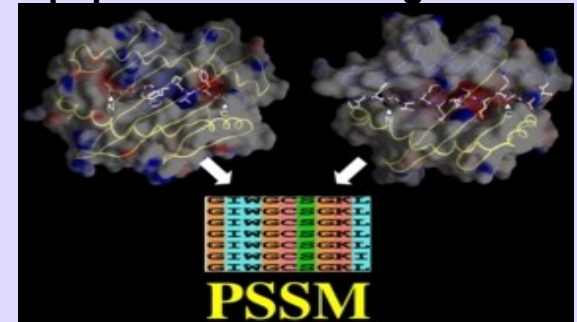


Computer tools for epitope discovery

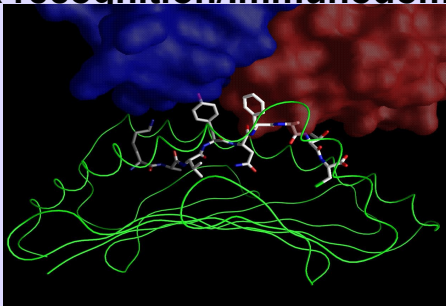
Antigen Processing



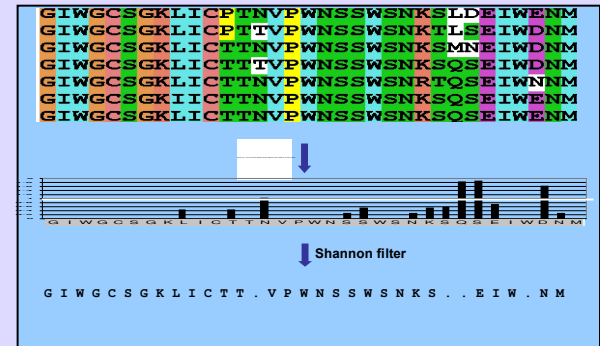
peptide-MHC binding



TCR recognition/Immunodominance




Variability parsing



RANKPEP:SERVIDOR PARA LA PREDICCION DE EPITOPOS

<http://bio.dfci.harvard.edu/Tools/rankpep.html>



Rankpep: prediction of binding peptides to Class I and Class II MHC molecules

Description
This server predicts peptide binders to MHC I and MHC II molecules from protein sequence/s or sequence alignments using Position Specific Scoring Matrices (PSSMs). In addition, it predicts those MHC I ligands whose C-terminal end is likely to be the result of proteasomal cleavage. A detailed explanation of the method can be found [here](#).

PSSM MHC I MHC II

H2-Db (mouse) [9mer]
H2-Db (mouse) [10mer]
H2-Db (mouse) [11mer]
H2-Dd (mouse) [9mer]
H2-Dd (mouse) [10mer]

HLA-DP4
HLA-DP9(DPA1*0201xDPB1*0901)
HLA-DPw4
HLA-DPw4(DPB1*0402)
HLA-DQ1

OR, UPLOAD YOUR PSSM no file selected

INPUT FASTA sequence/s CLUSTALW multiple sequence alignment

Replace example with your query
>A56881 PIR2 release 71.00
MWNLLHETDSAVATARRPRWLCAGALVLAGGFFLLGFLGFWFIKSSNEAT
NITPKHNMKAFLEDELKAENIKKFLYNFTQIPHLAGTEQNFQLAKQIQSQW
KEFGLDLSVELAHYDVLVLSYPNKTHPNYSIINEDGNEIFNTSLFEPPIPG

OR, UPLOAD SEQUENCES no file selected

BINDING THRESHOLD PERCENTAGE: 2% TOP NUMBER: 5

PROTEASOME CLEAVAGE FILTER: OFF LMPC: One
If Filter is ON only peptides predicted to be cleaved are shown

IMMUNODOMINANCE FILTER: OFF THRESHOLD: 59.4% sensitivity, 69.4% specificity
If Filter is ON only peptides to be immunodominant will be selected

ADVANCED OPTIONS


RESTRICT RESULTS BY MW
Lower Limit for Molecular Weight: 0.00
Upper Limit for Molecular Weight: 9999

VARIABILITY MASKING
Select Variability Threshold: 1
Value must range between 0.0 and 4.3

Citation:

- Reche PA, Glutting JP and Reinherz EL Prediction of MHC Class I Binding Peptides Using Profile Motifs. Human Immunology 63, 701-709 (2002).
- Reche PA, Glutting JP, Zhang H, Reinherz EL. Enhancement to the RANKPEP resource for the prediction of peptide binding to MHC molecules using profiles. 456:405-419 (2004)
- Manoj Bhasin, Ellis L. Reinherz and Pedro A. Reche. Modeling features of immunodominance into T-cell epitope identification. 2nd International Immunoinformatics Symposium. Boston University. March 7-9. 2005.

Questions, and suggestions to: [Pedro Reche](#)

Hits since June/2002 

Last updated: March/2004

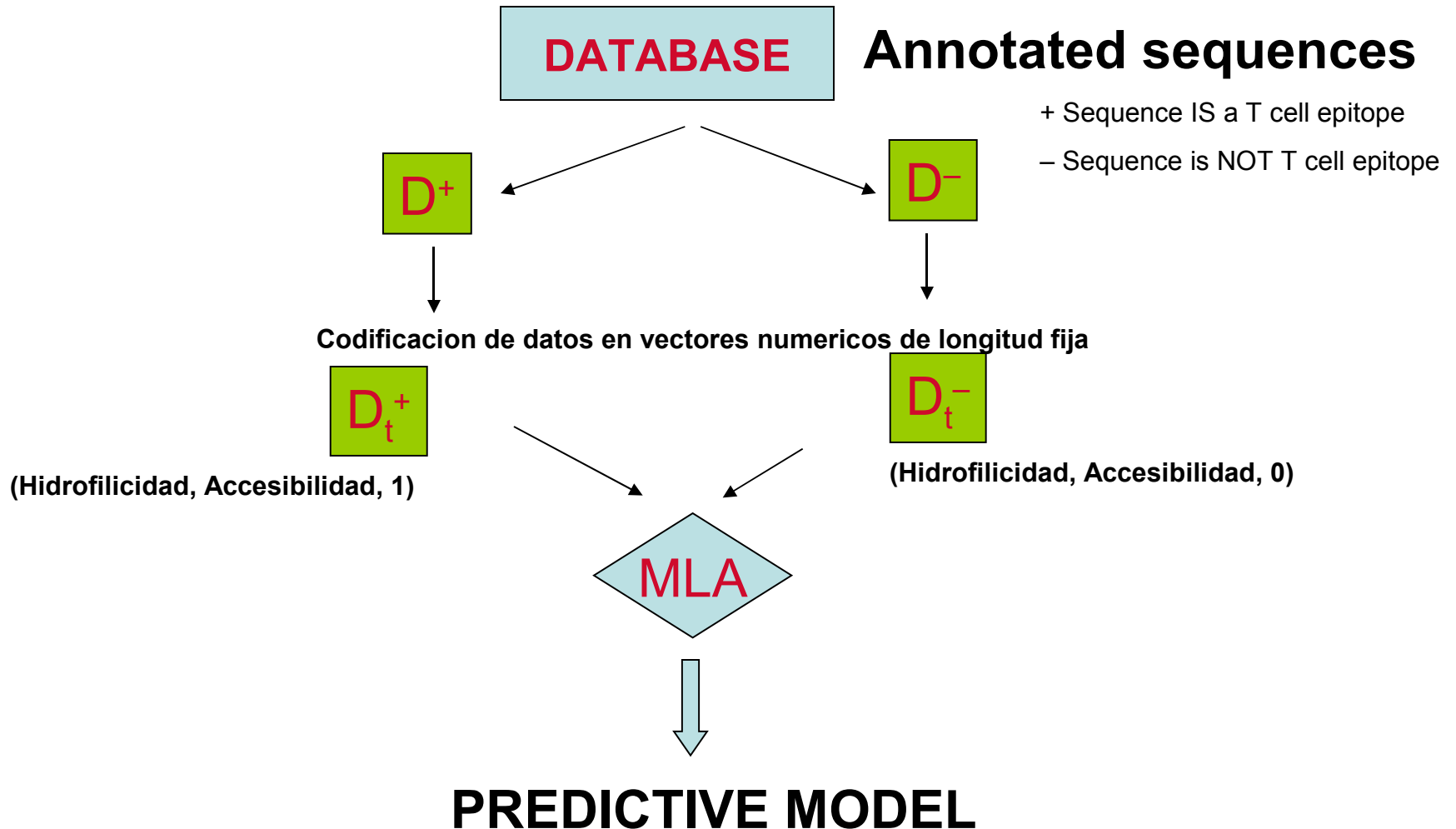
MIF Bioinformatics Molecular Immunology Foundation Saturday, 11/11/2006 15:18

- RANKPEP es la herramienta mas frecuentemente usada para al predicción de epitopos T (google y 400,000 hits)
- Posibilita la predicción de epitopos restringidos por moléculas MHC I y MHC II
- Proporciona la colección mas amplia de moléculas de MHC
- Predicción del corte por el proteosoma
- Es posible predecir cuales son los epitopos inmunodominantes
- Flexibilidad y versatilidad:
 - INPUT: Proteína/s o alineamientos de proteínas.
 - Posibilidad de limitar los resultados de acuerdo a su peso molecular
 - Posibilidad de identificar epítopos conservados

GENERAL PROCEDURE TO DERIVE DATA DRIVEN MODELS

1. **Data Collection.** Eg. Sequences of T cell epitopes
 - Databases Search
 - Text mining
- **Training.** Machine Learning Algorithms
- **Testing and Evaluation.**

GENERATION OF MLA-BASED PREDICTIVE MODELS



MLA: WEKA

- Support Vector Machine
- Decision Trees
- Artificial Neural Networks
- K-nearest neighbour learner

EVALUATION

- Select a parameter to quantify the performance of model (accuracy, MCC, etc)
- Evaluate performance through x-validation.
 - Train on a subset of the data and test in the remaining dataset

**Computer aided design of epitope
based vaccines: Computational
Vaccinology**

MOTIVACION/OBJETIVO

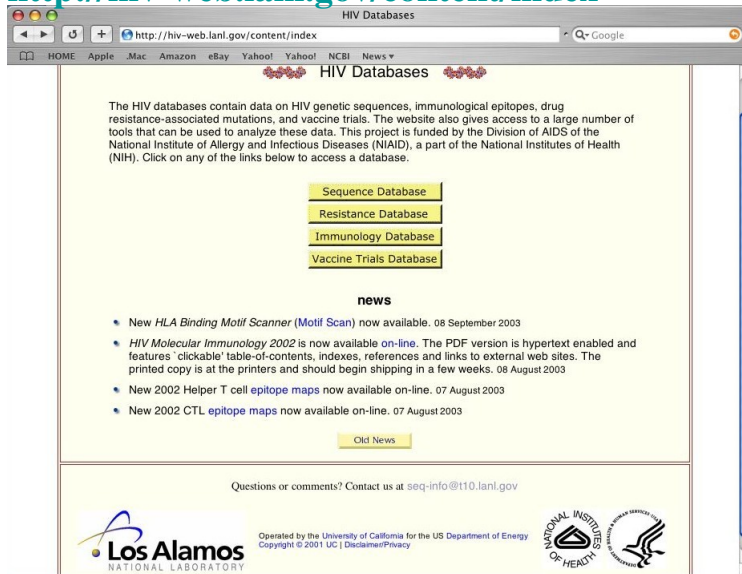
DESARROLLO DE UNA VACUNA DE EPITOPOS T CD8 FRENTE A HIV-1

Pedro A. Reche[§], et al (2006) Elicitation form virus-naive individuals of cytotoxic T lymphocytes directed against conserved HIV-1 epitopes. *Medical Immunology*, 5:1

KNOWN CTL EPITOPES IN HIV-1 INFECTED HUMANS

Los Alamos HIV database is depository of T-cell epitopes from HIV and SIV

<http://hiv-web.lanl.gov/content/index>



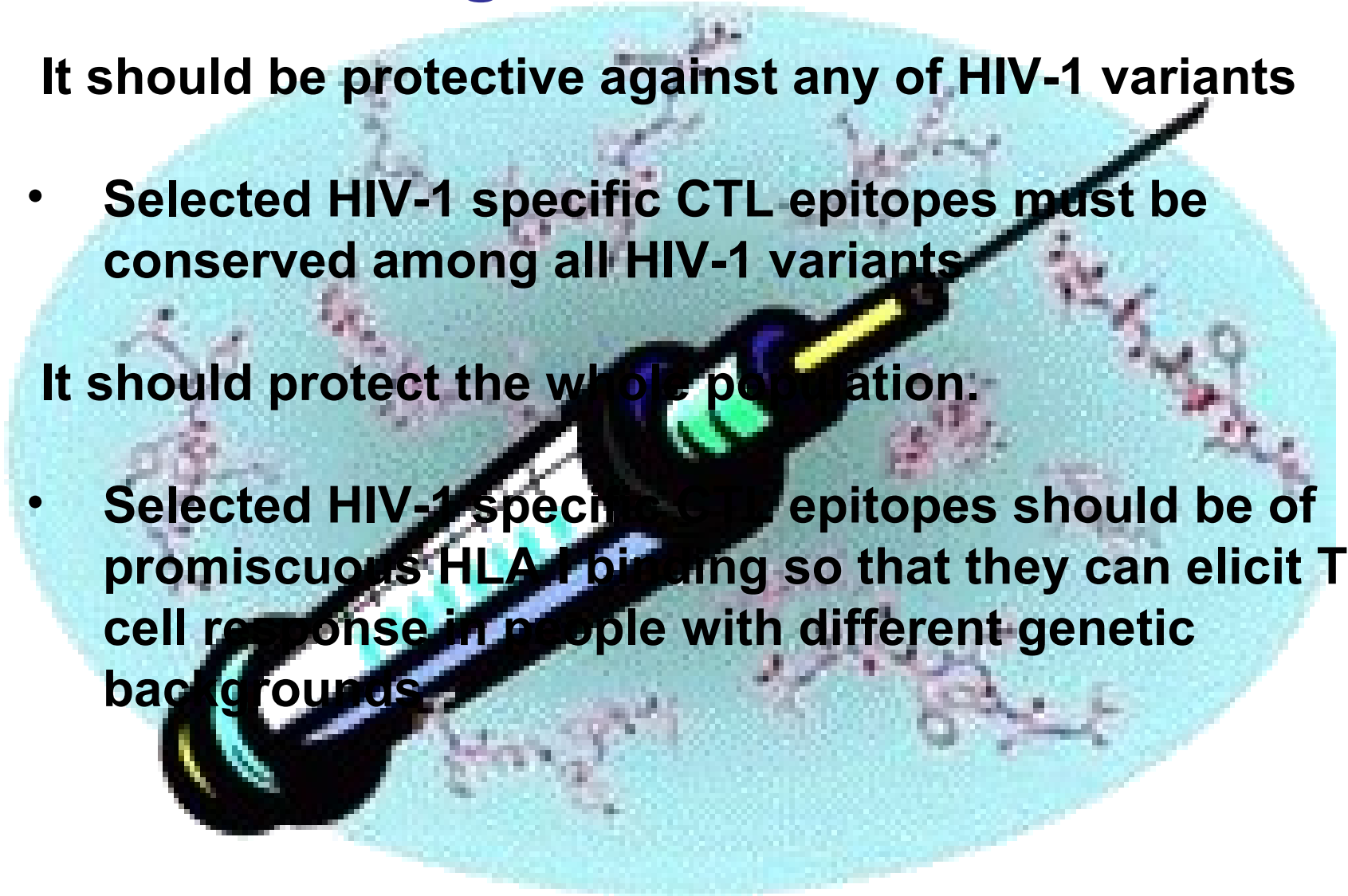
• CTL epitope example record

Displaying record number 1
HXB2 Location p17(18-26)
Author Location p17(18-26 IIIB)
Sequence KIRLRPGGK
Species(HLA) human (A3)
Immunogen HIV-1 infection
Keywords
Notes
References

• 1567 CTL epitope records => 592 unique seq =>195 9mers
Immunogen HIV-1 infection
Species(HLA) human

• CTL epitopes from HIV infected patients infected could be used as the basis of a vaccine against HIV1. **How?**

Criteria for a CTL epitope vaccine against HIV-1

1. It should be protective against any of HIV-1 variants
 - Selected HIV-1 specific CTL epitopes must be conserved among all HIV-1 variants
 2. It should protect the whole population:
 - Selected HIV-1 specific CTL epitopes should be of promiscuous HLA binding so that they can elicit T cell response in people with different genetic backgrounds
- 

HIV-1 SEQUENCES

Table 1

HIV-1 protein sequences used in this study

GENE	Sequences*																		
	Total	A-A1	B	C	D	F1	F2	G	H	K	O	01	02	04	06	10	11	12	CPZ
GAG	207	10	37	28	5	7	2	4	3	3	5	14	8	3	4	3	4	3	5
POL	204	11	34	26	5	4	3	4	3	2	5	14	12	3	4	3	5	3	5
ENV	408	24	128	51	17	5	5	9	3	2	10	33	16	3	5	3	6	3	5
VIF	427	28	190	40	13	5	5	5	3	2	24	15	12	3	4	3	4	12	5
TAT	219	11	39	29	7	6	3	4	3	2	4	14	14	3	6	3	4	3	5
REV	231	13	39	38	10	6	3	4	3	2	4	17	12	3	4	3	4	3	5
VPU/VPX	333	14	87	50	18	6	3	6	3	2	15	25	14	3	4	3	4	3	5
VPR	368	9	142	44	7	4	3	5	3	2	25	16	12	3	4	3	4	3	5
NEF	449	12	282	44	4	4	3	7	5	2	4	31	12	3	4	3	11	3	4

*HIV-1 sequences were obtained from the HIV database for the indicated clade categories.

Selection of Conserved CTL

HIV1

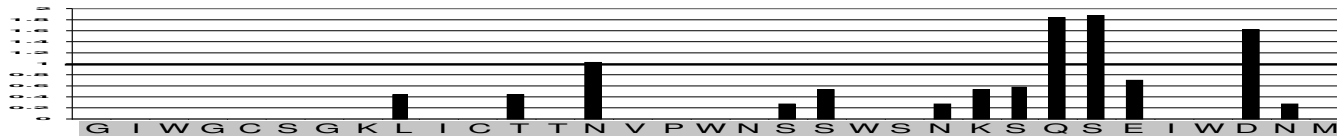
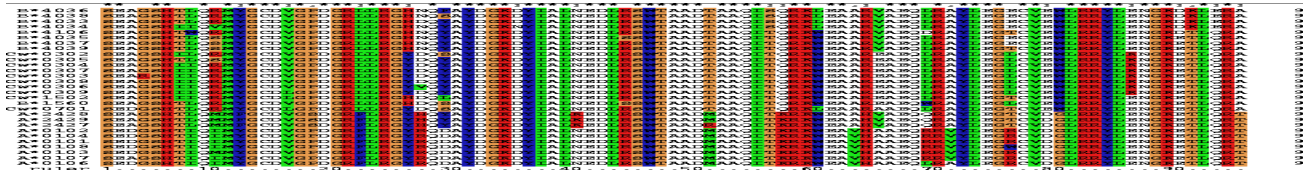
GAG
POL
ENV
VIF
TAT
REV
VPU/VPX
VPR
NEF

Clustalw
→

GAG.aln
POL.aln
ENV.aln
VIF.aln
TAT.aln
REV.aln
VPU.aln
VPR.aln
NEF.aln

Mask variability
→

GAG.seq
POL.seq
ENV.seq
VIF.seq
TAT.seq
REV.seq
VPU.seq
VPR.seq
NEF.seq



H=1



Shannon filter

G I W G C S G K L I C T T . V P W N S S W S N K S . . E I W . N M

Only HIV-1 specific CTL epitopes that match the consensus sequences are considered for further analysis

Conserved HIV1 CTL Epitopes

H < 1			
HIV CTL	SOURCE	POS	ALLELES (HIVDB)
SPRTLNAWV	p24	16-24	B0702
AVFIHNFKR	Integrase	179-187	A0301
MAVFIHNFK	Integrase	178-186	A0301
TLFCASDAK	gp160	51-59	A0301
FPVRPQVPL	Nef	68-76	B3501
RAMASDFNL	Integrase	20-28	A0201
KLTPLCVTL	gp160	121-129	A0201
TLNAWVKVI	p24	19-27	A0201
VIYQYRDDL	RT	179-187	A0201
LVGPTPVNI	Protease	76-84	A0201
TVLDVGDAY	RT	107-115	B3501
PLVKLWYQL	RT	421-429	A0201
TLNFPISPI	POL	POL	A0201
NTPVFAIKK	RT	57-65	A0301
SEGATPQDL	p24	44-52	
EKEGKISKI	RT	42-50	B5101
LLWKGEHAV	Integrase	241-249	A0201
KLVGKLNWA	RT	259-267	A0201
LTFGWCFKL	Nef	137-145	A0201
YQYMDLYV	RT	181-189	A0201
GPKVKQWPL	RT	18-26	B0801
WASRELERF	p17	36-44	B3501
RAIEAQOHL	gp160	557-565	B5101 B1501 C0304
GLNKIVRMY	p24	137-145	B1501
KEKGGLEGL	Nef	92-100	B4002
YFPDQNYT	Nef	120-128	A1 B3701 B5701
WYIKIFIMI	gp160	680-688	A2402
YVDRFFKTL	p24	164-172	A2601
FVNTPLVK	RT	416-424	A1101
DRFFKTLRA	p24	166-174	B1402
KIQNFRVYY	Integrase	219-227	A3002
KLNWASQIY	RT	263-271	A3002
QGWKGSPAI	RT	151-159	B5101
IRLRPGGKK	p17	19-27	B2705
DLSHFLKEK	Nef	86-94	A0301
KIRLRPGGK	p17	18-26	A0301 B0301
GIPHPAGLK	RT	93-101	A0301
MTKILEPFR	RT	164-172	A0301
AETFYVDGA	RT	437-445	B4501
EKAFSPEV	p24	28-36	B4415
CRAPRKKGK	p2p7p1p6	42-50	B1402
ITLWQRPLV	Protease	03-11	

↔
Conser.

H < 0.5			
CTL	SOURCE	POS	ALLELES (HIVDB)
SPRTLNAWV	p24	16-24	B0702 B7
AVFIHNFKR	Integrase	179-187	A0301
MAVFIHNFK	Integrase	178-186	A0301
LVGPTPVNI	Protease	76-84	A0201 A2
TVLDVGDAY	RT	107-115	B3501 B35
GLNKIVRMY	p24	137-145	B1501
SEGATPQDL	p24	44-52	B60
PLVKLWYQL	RT	421-429	A0201
LLWKGEHAV	Integrase	241-249	A0201
FVNTPLVK	RT	416-424	A1101
KLVGKLNWA	RT	259-267	A0201
YQYMDLYV	RT	181-189	A0201
QGWKGSPAI	RT	151-159	B5101 B62
GIPHPAGLK	RT	93-101	A0301
LTFGWCFKL	Nef	137-145	A0201
TLNFPISPI	POL		A2
KLNWASQIY	RT	263-271	A3002 A30

Total: 17 peptides



Experimental HLA binding from HIV database

Total: 42 peptides

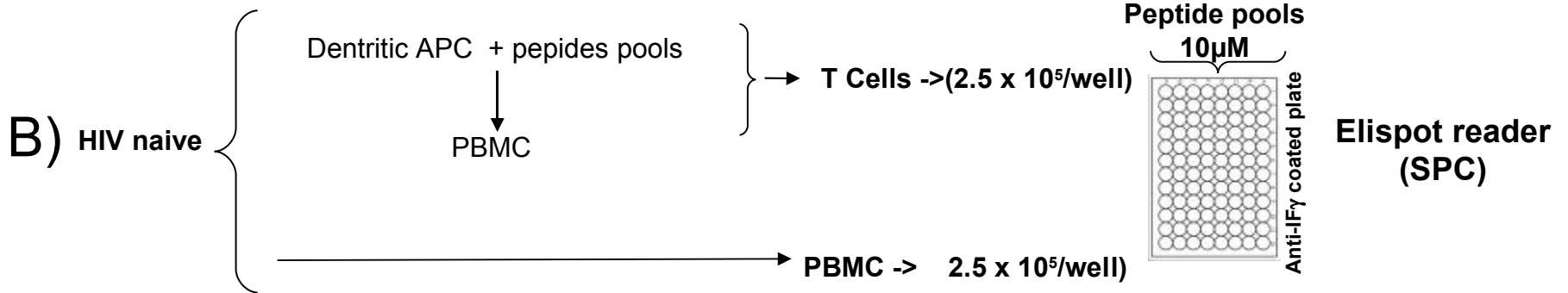
Extended HLA binding Profile of HIV1 specific CTL Epitopes

HIV CTL	SOURCE	POS	ALLELES (HIVDB)	ALLELES (HIVDB + PREDICTED)
KLNWASQIY	RT	263-271	A3002	A1 A3002
TVLDVGDAY	RT	107-115	B3501	B1501 B3501 B5701 C0304
RAMASDFNL	Integrase	20-28	A0201	A0201 B2709 C0304
KIRLRPGGK	p17	18-26	A0301 B0301	A0301 B0301
KEKGGLEGL	Nef	92-100	B4002	B4002 B4402 B4403
PLVKLWYQL	RT	421-429	A0201	A0201 A0202 A0203
YVDRFFKTL	p24	164-172	A2601	A0203 A0204 A0207 A2601 B3801
SEGATPQDL	p24	44-52		A2902 B39011 B4002 B4402 B4403
LVGPTPVNI	Protease	76-84	A0201	A0201 A0202 A0205 A0209 B1501 B1516
GLNKIVRMV	p24	137-145	B1501	A0203 A1 B1501
GPVKQWPL	RT	18-26	B0801	B0702 B0801 B3501 B8
DRFFKTLRA	p24	166-174	B1402	B1402 B2701 B2702 B2703 B2704 B2705 B2709
TLNAWKVI	p24	19-27	A0201	A0201 A0202 A0203 A0204 A0206
QGWKGSFAI	RT	151-159	B5101	B5101
KLVGKLNWA	RT	259-267	A0201	A0201
LTPGWCFKL	Nef	137-145	A0201	A0201
YFPDQNYT	Nef	120-128	A1 B3701 B5701	A1 B3701 B5701
NTPVFAIKK	RT	57-65	A0301	A0301 A6601 C0102
AETFYVDGA	RT	437-445	B4501	B4501
EKEGKISKI	RT	42-50	B5101	B2701 B3801 B39011 B3909 B4402 B5101 B8
DLSHFLKEK	Nef	86-94	A0301	A0301 A6601
SPRTLNAWV	p24	16-24	B0702	B0702 B3501 B5101 B5102 B5103 B5301 B5401 B5502
AVFIHNFKR	Integrase	179-187	A0301	A0301 A1101 A3101 A3301 A6601 A6801
VIYQYMDL	RT	179-187	A0201	A0201 A0205 A0207 A0214
MTKILEPFR	RT	164-172	A0301	A0301
MAVFIHNFK	Integrase	178-186	A0301	A0301
ITLWQRPLV	Protease	03-11		A6802
FVNTPLLVK	RT	416-424	A1101	A1101
TLFCASDAK	gp160	51-59	A0301	A0301 A1101 A3101 A3301 A6801
FPVRPQVPL	Nef	68-76	B3501	A2902 B0702 B3501 B5101 B5102 B5103 B5301 B5401
IRLRPGGKK	p17	19-27	B2705	B2705
WASRELERF	p17	36-44	B3501	B3501 B5801
TLNFPISPI	A2	A2		A0201 A0207
LLWKGEGAV	Integrase	241-249	A0201	A0201 A0204 A0205 A0209
KLTPLCVTL	gp160	121-129	A0201	A0201 A0202 A0203 A0206 A0209 B2709
RAIEAQOHL	gp160	557-565	B5101 B1501 C0304	B1501 B1517 B5101 C0304
EEKAFSPEV	p24	28-36	B4415	B4415
GIPHPAGLK	RT	93-101	A0301	A0301
YQYMDLLYV	RT	181-189	A0201	A0201
WYIKIFIMI	gp160	680-688	A2402	A0203 A0206 A2402
CRAPRKKGC	p2p7p1p6	42-50	B1402	B1402
KIQNFRVYY	Integrase	219-227	A3002	A1 A3002

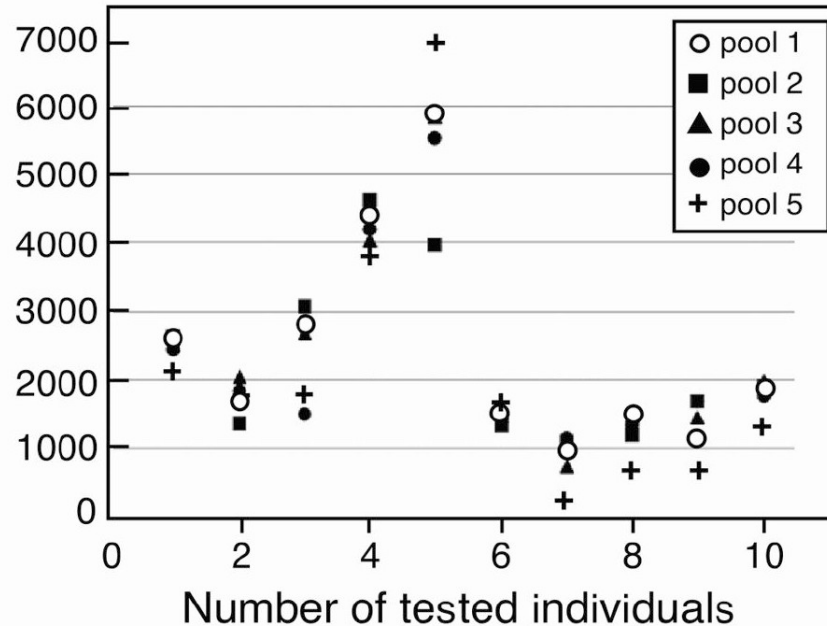
Apply an algorithm to identify combinations of epitopes providing a population coverage of 95%

A minimum of 5 peptides are required to cover the whole population

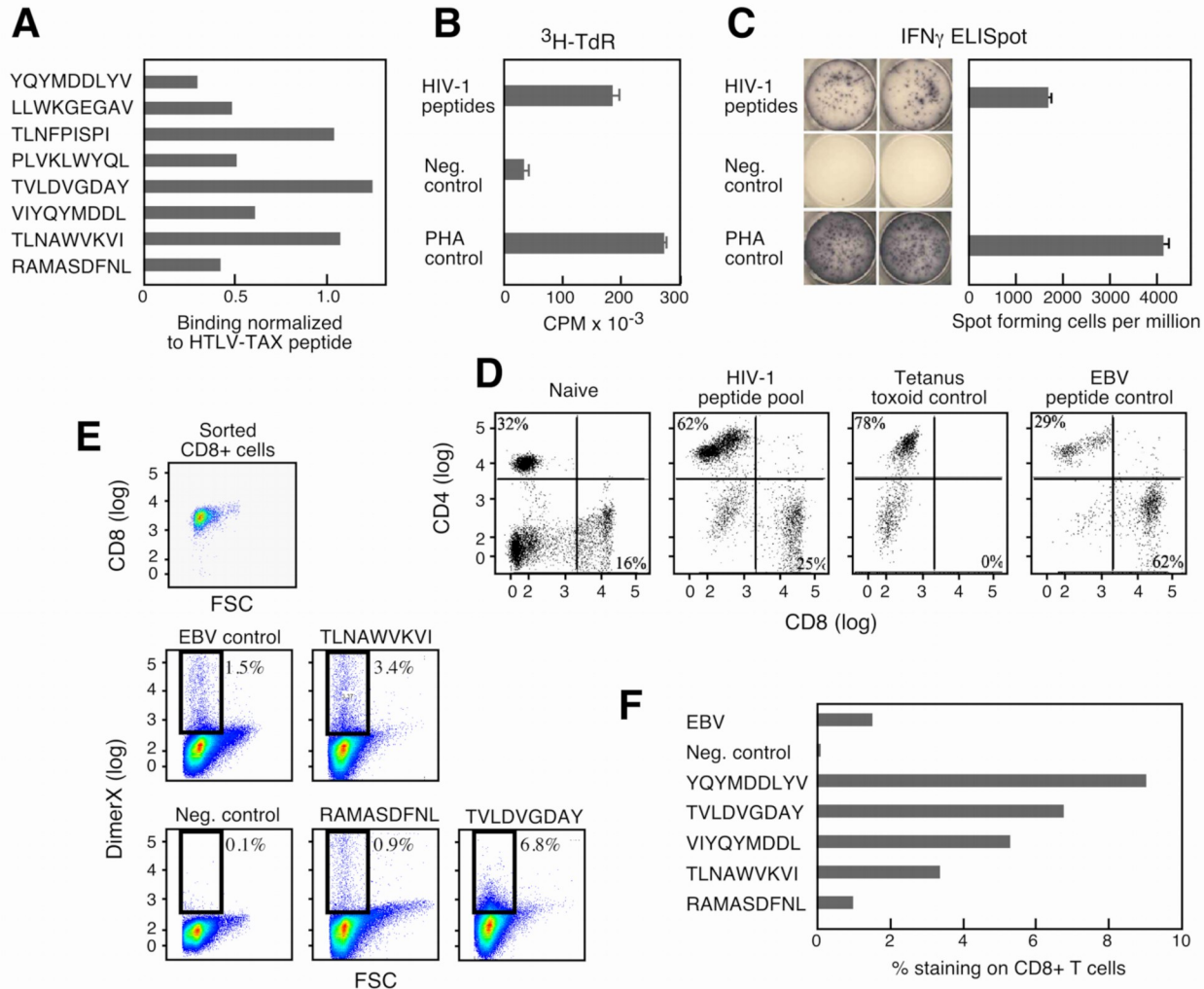
Responses of stimulated T cells from naïve people to CTL pools



B HIV-naïve normal donors primed in vitro

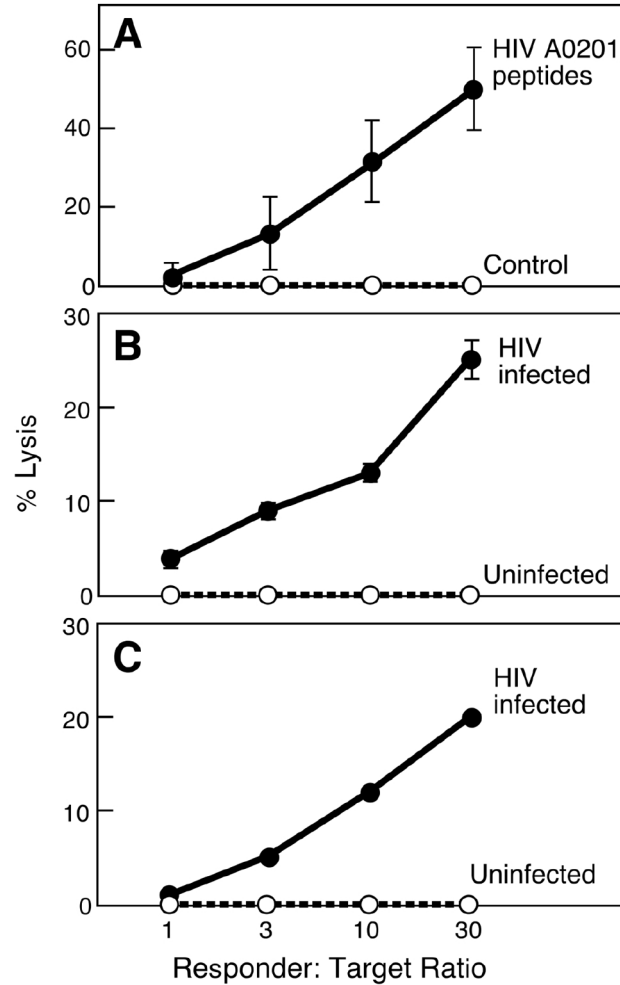


Characterization of HLA-A0201 restricted HIV-1 peptide-specific T cell lines from uninfected individuals



Cytotoxic activity of HIV-1 peptide-specific A0201-restricted T cells from naive donors

T2 cells + peptide +
A0201-restricted HIV-1 specific T cell

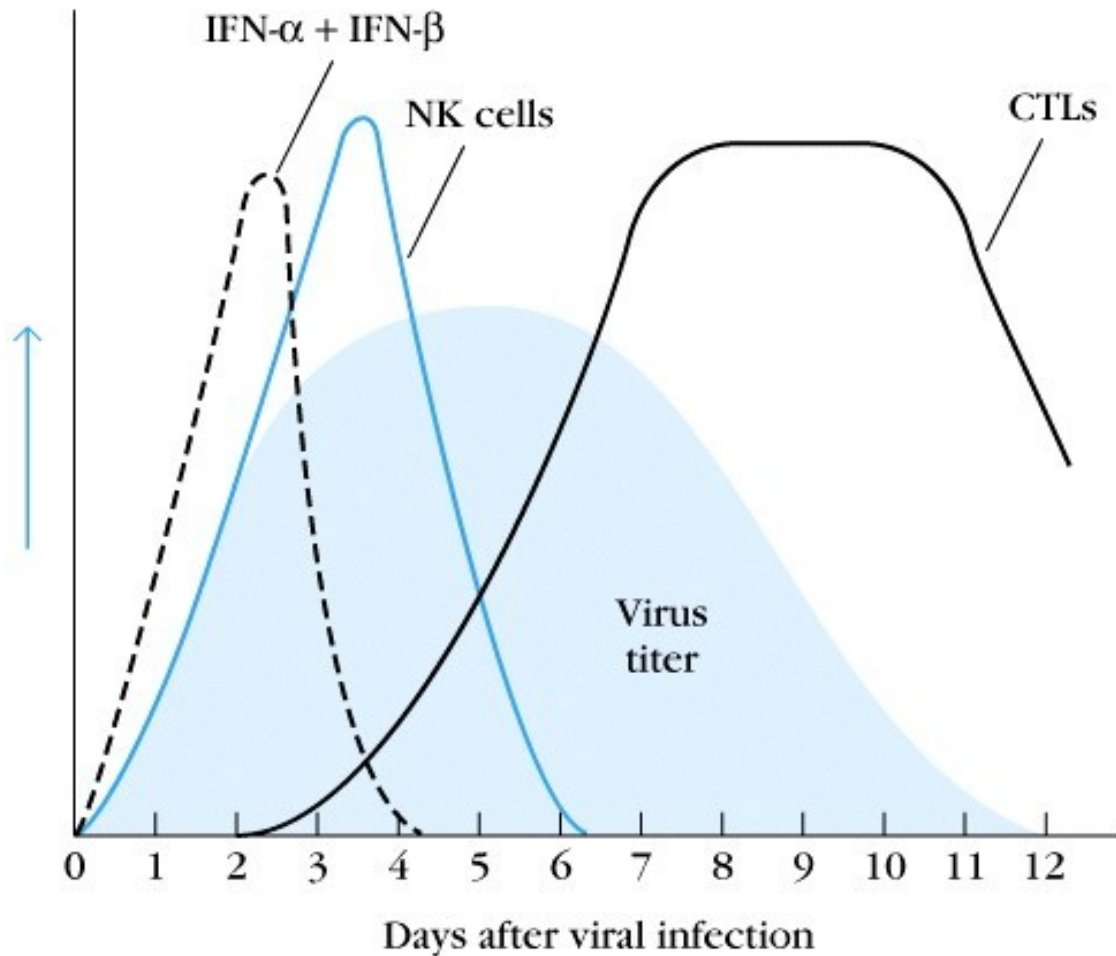


T1 cells infected/uninfected with HIV-III B

Conclusions

- **Naïve individual have the potential to recognize invariant HIV-1 specific T cell epitopes**
- **Chronic infections and diseases such as cancer may result in the impairment of T cell immune response and therefore epitope vaccine optimization should be carried out from data obtained in naïve healthy patients**
- **Chronic infections and diseases escape the immune system due the emergence of escape mutants that are no longer the target of the immune system and/or by directing the immune response towards antigenic determinants that results in futile pattern of immunodominance. These two issues can be better addressed using epitope-based vaccines that allow to focus the immune response towards the desired regions (strictly conserved epitopes)**

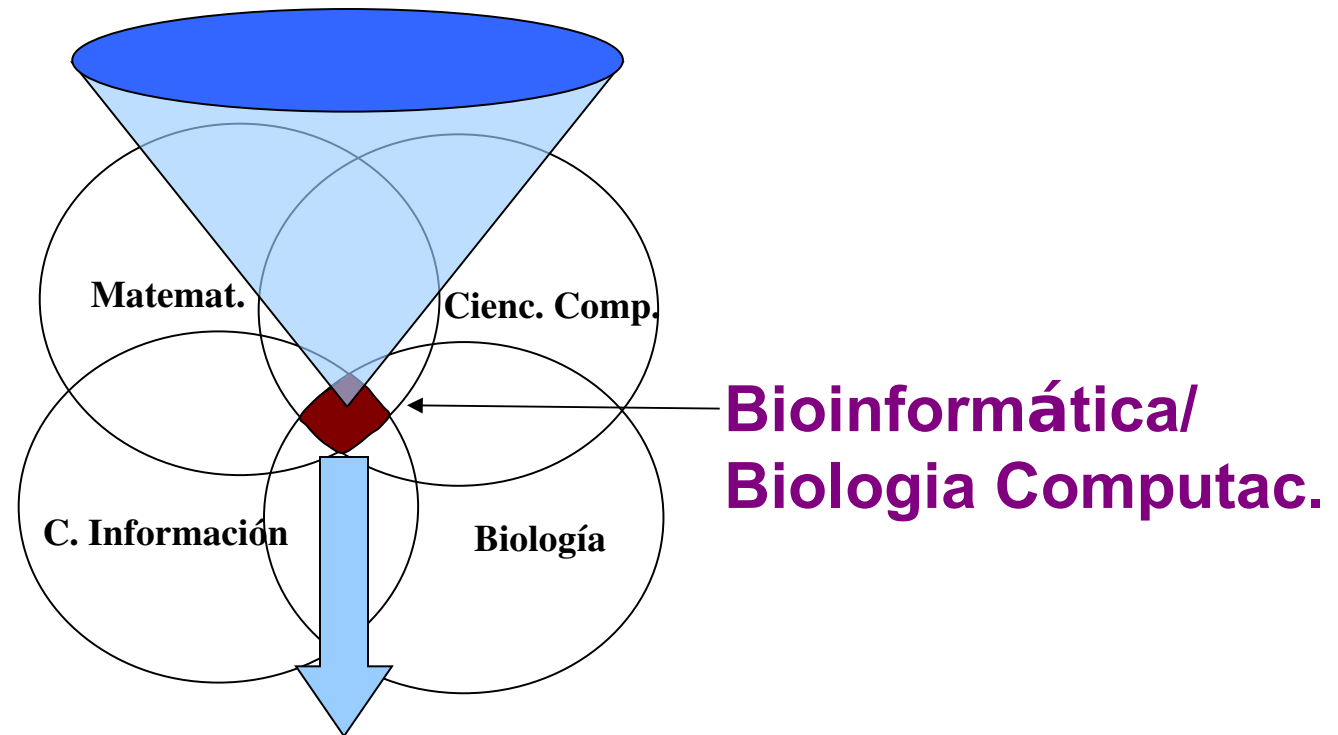
Innate responses to virus infection



BIOINFORMATICA

- La Bioinformática podría definirse como el campo de la ciencia en el que las Matemáticas, las Ciencias Computacionales y de la Información se encuentran con la Biología para crear una disciplina nueva.

DATOS Y INFORMACION/HIPOTESIS



Bases de Datos, herramientas de gerencia
Modelos Computacionales para el análisis e interpretación de datos
Conocimiento Biológico