Inserm
U599 Cancer Institute of Marseille

INSTITUT
CENTRE RÉGIONAL DE LUTTE CONTRE LE CANCER
PAOLI-CALMETTES
PROVENCE-ALPES-CÔTE D'AZUR

UNIVERSITÉ DE LA MEDITERRANÉE AIX-MARSEILLE II

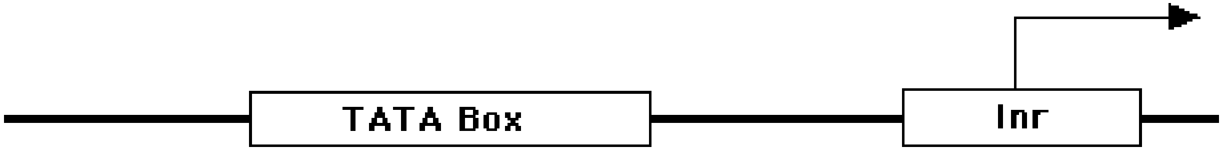
Analysis of Gene Regulatory Regions using Comparative Genomics: From the wet lab to computer and back

Jean Imbert, INSERM U599
Tel: 04 91 75 84 04 - imbert@marseille.inserm.fr
http://u119.marseille.inserm.fr/ji.html

icim
Institut de Cancérologie
et d 'Immunologie de Marseille

INSERM Workshops 2004 n°152:
Comparative genomics of vertebrate: concepts and bioinformatic tools
May 27-28, 2004
La Londe – Les Maures, France

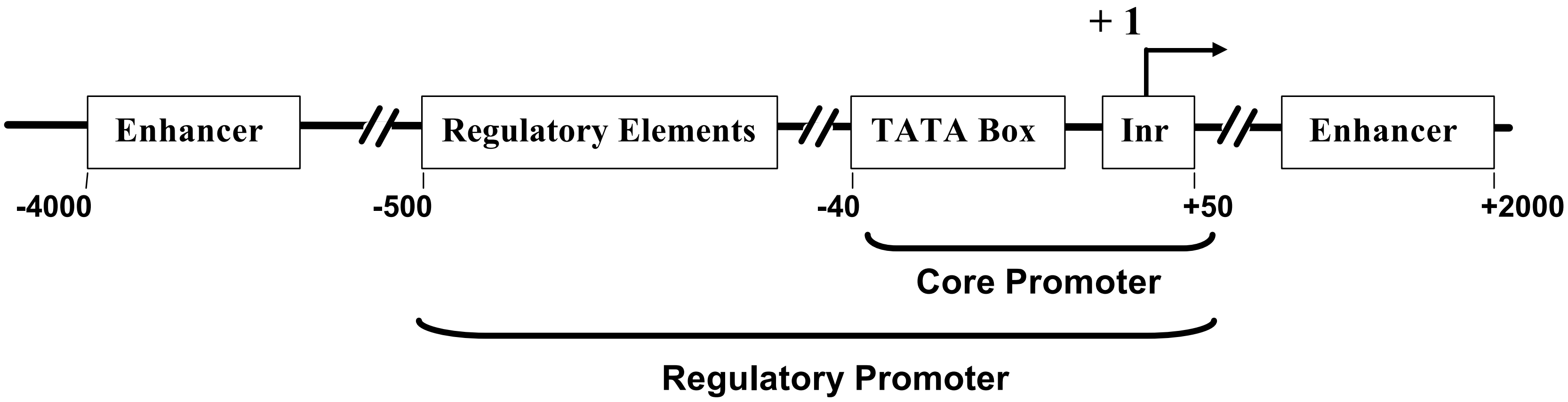
CORE PROMOTER



TATA Box: TATAAAA (about 25 base pairs upstream of the start point)

Initiator (Inr): PyPyCAPyPyPyPyPy

MODEL OF TYPICAL GENE PROMOTER AND REGULATORY REGIONS



Eukaryotic Promoter Classes

Pol I	< 1%	3/4 rRNAs (28S, 18S, 5.8S)
Pol II with TATA-box	> 70%	mRNAs
Pol II without TATA-box	~ 20%	
Pol III internal	~ 5%	Small RNAs tRNAs 5S RNAs
Pol III upstream with TATA-box	< 1%	
Pol III upstream without TATA-box	< 1%	

Components of Eukaryotic Promoters and Regulatory Regions

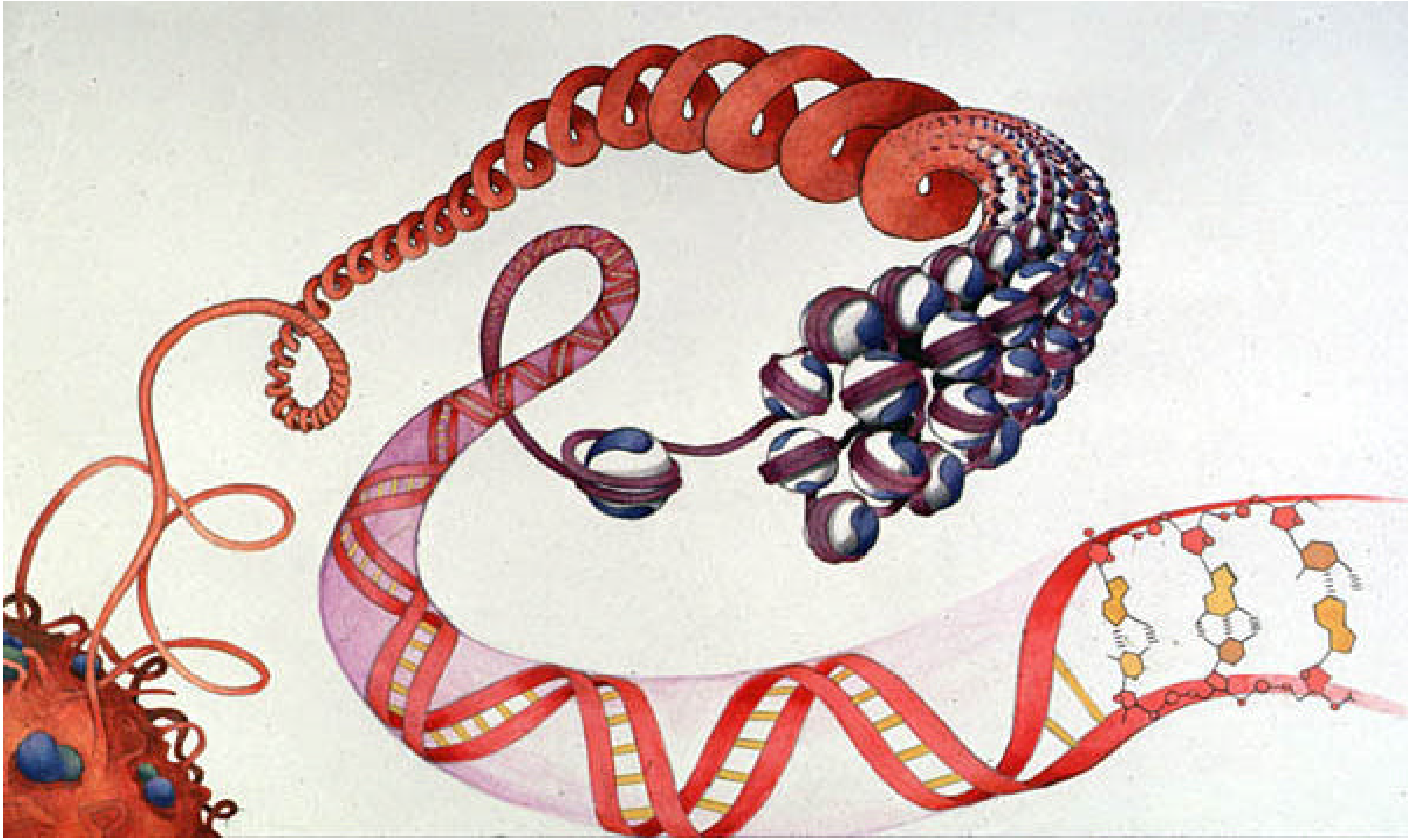
• Site selector elements	TATA-box, Initiator
• Common upstream elements	CCAAT-box, GC-box
• Regulatory elements	HSE, SRE, GRE, etc.

- Enhancers / Silencers
- Locus control regions (LCRs)
- Scaffold / Matrix attachment sites (SARs / MARs)

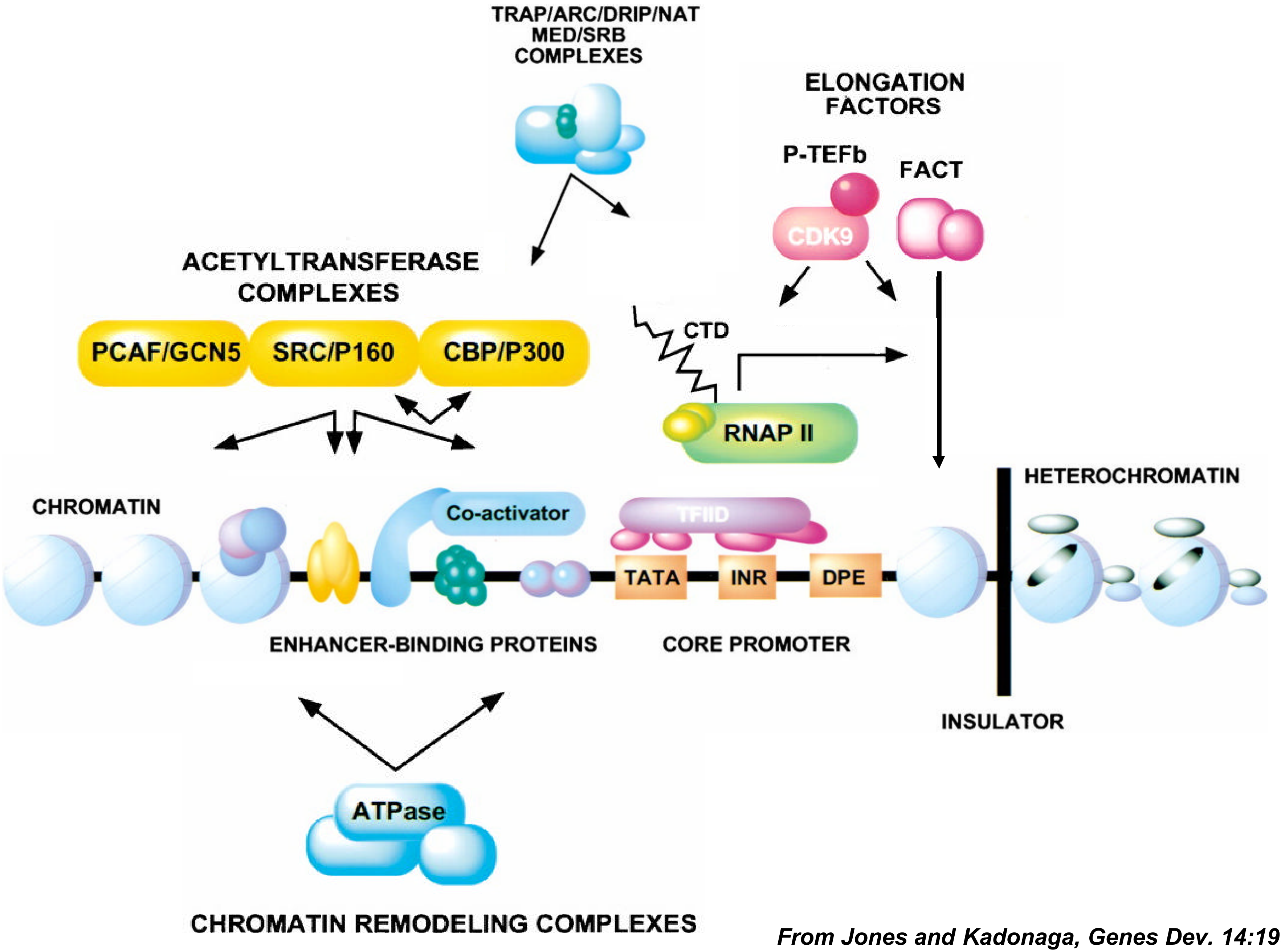
- CpG islands

Promoter Regulatory Elements: Features and Facts

- Degenerate sequence motifs
- Length: 6 to 20 bp
- Low complexity (8-12 bits)
- Binding sites of transcription factors
- Excess of binding sites over binding proteins in the nucleus
- Most in vitro binding sites not functional in vivo
- Some in vivo binding sites also not functional
- Regulatory potentials depends on cooperative effects between multiple elements



From G. Almouzni, Institut Curie, Paris



From Jones and Kadonaga, Genes Dev. 14:1992-1996, 2000.

TOOLS FOR THE IDENTIFICATION OF REGULATORY SEQUENCES AND THEIR COGNATE TRANSCRIPTION FACTORS

1. Identification and characterization of regulatory sequences

- Gene reporter assays: transient or stable transfection (CAT, Luciférase, SEAP, GFP, b-gal, etc.)
- *In vitro* transcription assay
- Sequencing, database mining (web: TESS, Euk. Pr. Database, TRANSFAC, MATINSPECTOR, TFSEARCH, etc.)
- Animal or cellular models: enhancer trap, enhancer knock-in, minichromosomes, etc.

2. Identification and characterization of specific transcription factors (TFs)

- EMSA and sequels: UV-crosslinking, pull-down assay using biotynilated oligonucleotide
- Footprint detection:
  - * *in vitro :* nucleases (DNAse I hypersensitivity, S1 nuclease, Mmase) or chemical compounds
  - * *in vivo* : genomic footprinting, Chromatin ImmunoPrecipitation (ChIP) and sequels

3. TFs physical and functional interactions

- Transfection and biochemistry
- ChIP-on-chip

Inserm
U599 Cancer Institute of Marseille

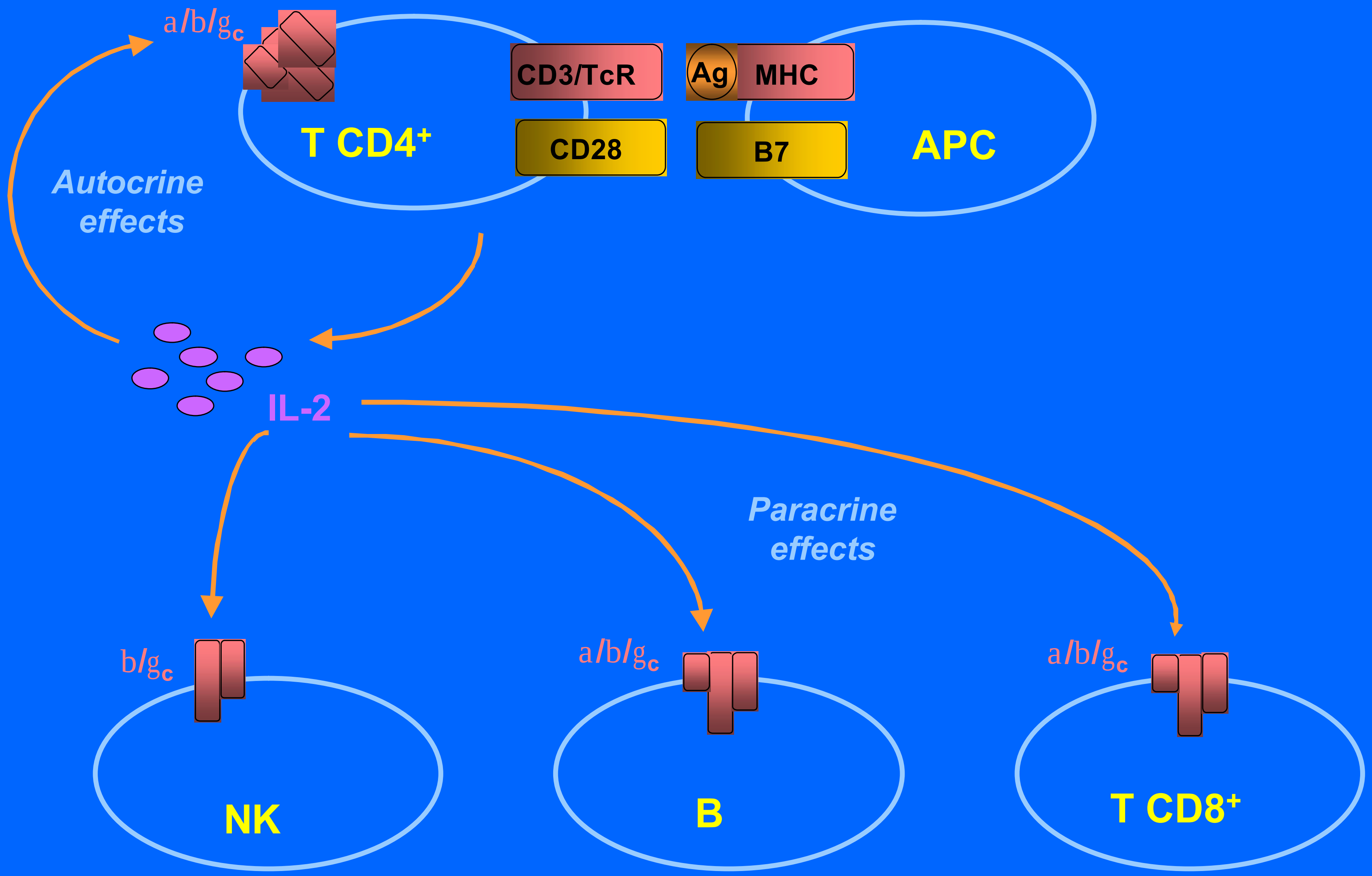
INSTITUT
CENTRE RÉGIONAL DE LUTTE CONTRE LE CANCER
PAOLI-CALMETTES
PROVENCE-ALPES-CÔTE D'AZUR

UNIVERSITÉ DE LA MÉDITERRANÉE
AIX-MARSEILLE II

Control of CD25/IL-2Ra gene transcription

icim
Institut de Cancérologie
et d 'Immunologie de Marseille

Primary Activation



ROLE OF THE IL-2Ra CHAIN

- SOLE IL-2 SPECIFIC CHAIN

- *DE NOVO* EXPRESSION CREATES AN HIGH AFFINITY IL-2 RECEPTOR (Kd=10^{-11}) IN ASSOCIATION WITH THE INTERMEDIATE AFFINITY RECEPTOR b/g$_c$ (Kd=10^{-9}) SHARED WITH IL-4R, -7R, -9R AND -15R

- IL-2 $^{-/-}$ MICE ~ IL-2Ra $^{-/-}$ MICE

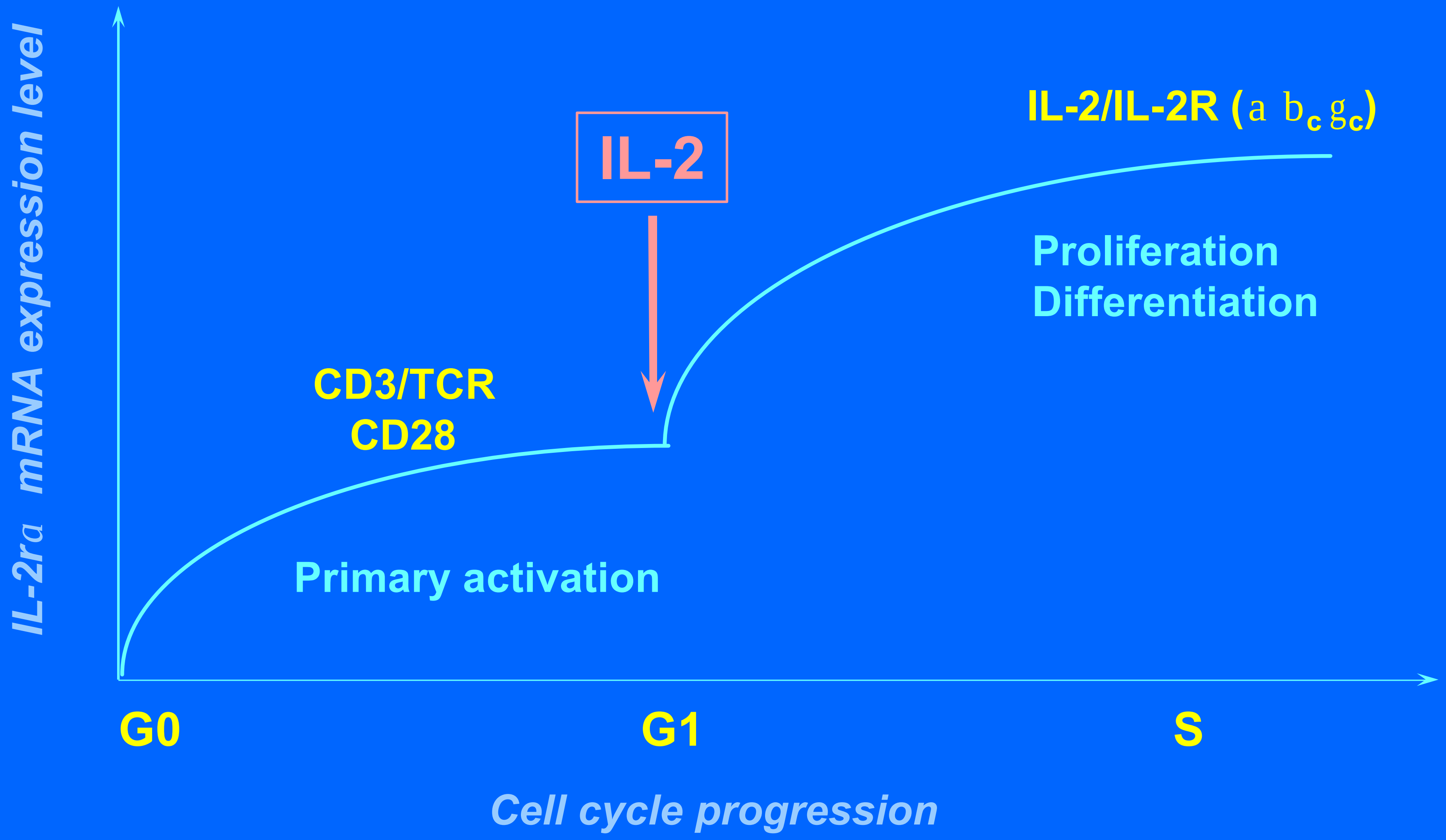
THE HIGH AFFINITY IL-2 RECEPTOR IS ESSENTIAL TO TRIGGER AND SUSTAIN AN EFFICIENT RESPONSE TO THE LIMITED AMOUNT OF SECRETED IL-2 AVAILABLE IN PHYSIOLOGICAL CONDITIONS

IL-2Ra CHAIN EXPRESSION IS CONTROLLED A TWO MAIN LEVELS:

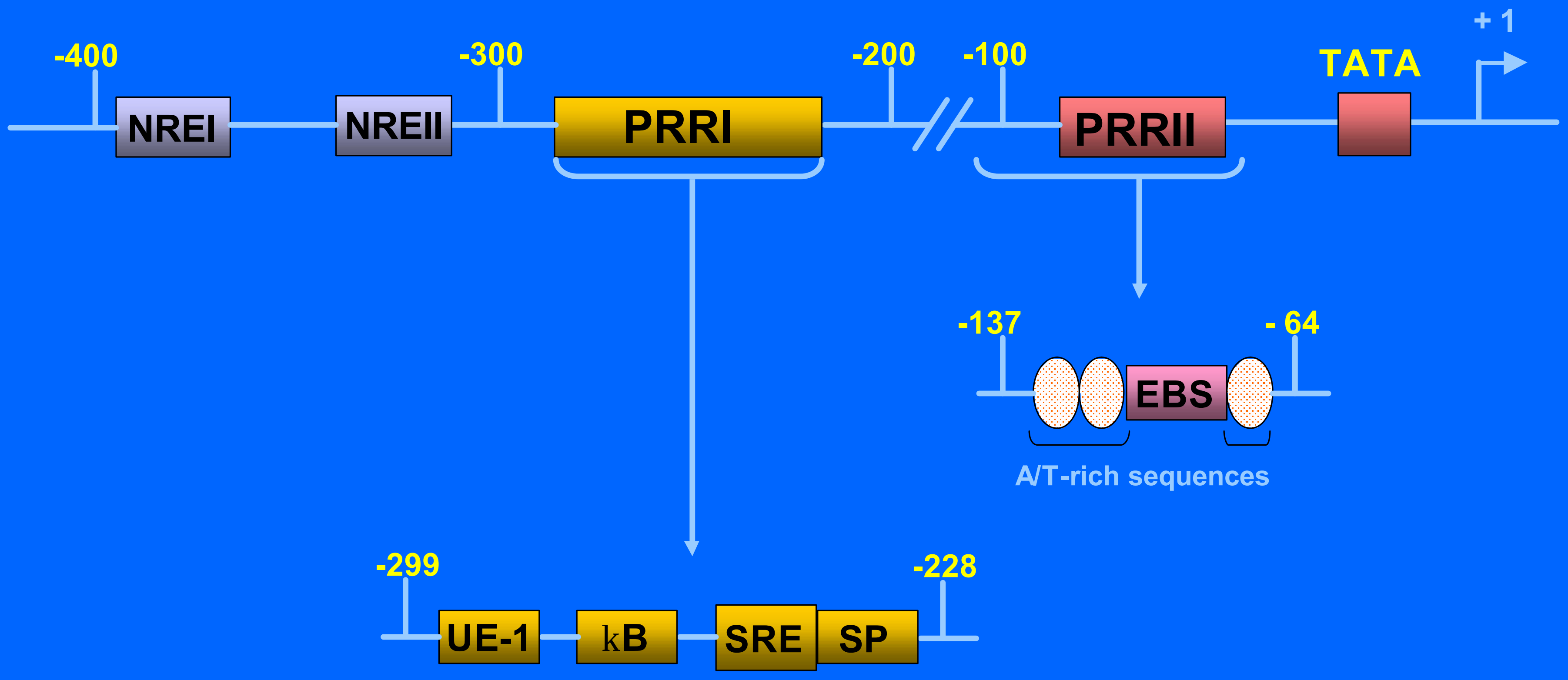
- TRANSCRIPTIONAL

- POST-TRANSCRIPTIONAL
  (mRNA stabilization)

CD25/IL-2Ra GENE TRANSCRIPTION DURING T CELL ACTIVATION

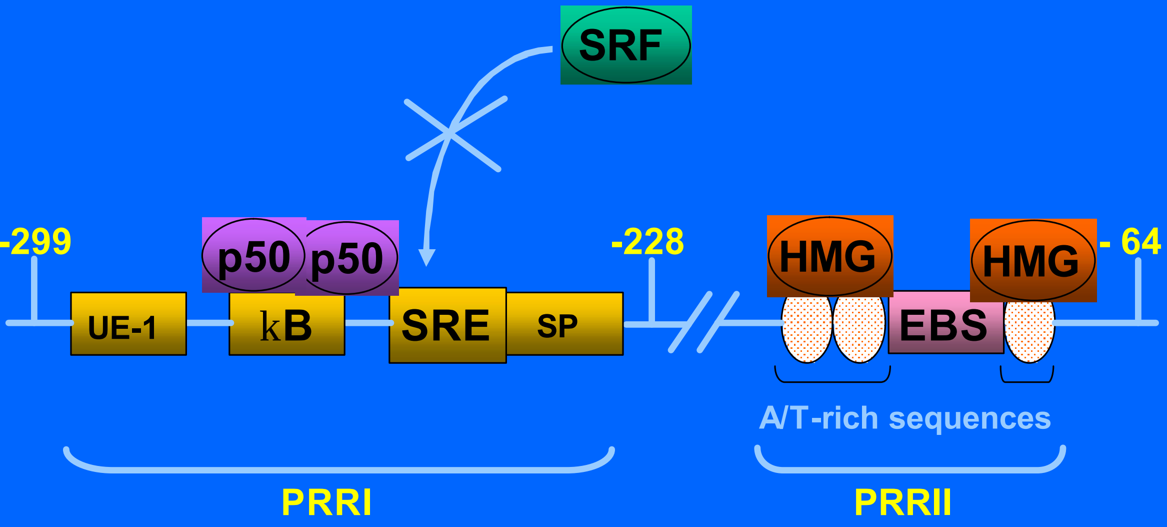


PROXIMAL REGULATORY REGIONS OF THE HUMAN CD25/IL-2Ra GENE



PRR: POSITIVE REGULATORY REGION

RESTING T LYMPHOCYTE



PRRI : Inactive NF-kB p50/p50 homodimers binding

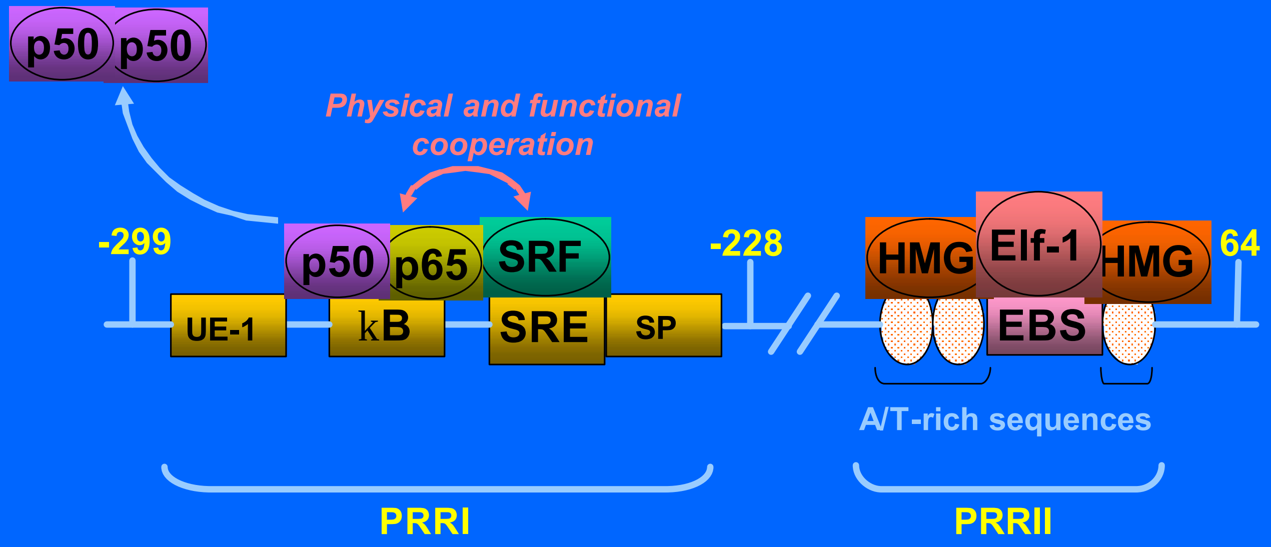
PRRII : HMGI(Y) constitutive binding

Algarte, M., Lecine, P., Costello, R., Plet, A., Olive, D. and Imbert, J. In vivo regulation of interleukin-2 receptor alpha chain gene transcription by the coordinated binding of constitutive and inducible factors in human primary T-cells. EMBO J. 14: 5060-5072; 1995.

Primary activation

↓

ACTIVATED T LYMPHOCYTE



PRRI
(induction)

- Active NF-kB p50/p65 or p65/c-Rel heterodimer binding
- SRF binding
- Physical and functional intercation SRF / NF-kB

Ballard et al., Science, 1988; Toledano et al., PNAS, 1990; Costello et al., CGD, 1993; Algarte et al., EMBO J, 1995; John et al., MCB, 1995

PRRII
(T cell specificty)

- Elf-1 binding

John et al., MCB, 1995

Homo Sapiens versus Mus Musculus PRRI/II conservation

Local identities (aligment n°6)

```
8724-8798 <--> 11629-11703 83% (75 nt)
8800-8886 <--> 11704-11790 76% (87 nt)
8887-8911 <--> 11792-11816 96% (25 nt)
8913-8946 <--> 11817-11850 97% (34 nt)
8954-8979 <--> 11851-11876 100% (26 nt)
8982-8999 <--> 11877-11894 83% (18 nt)
9003-9006 <--> 11895-11898 100% (4 nt)
9007-9015 <--> 11901-11909 78% (9 nt)
9017-9079 <--> 11910-11972 70% (63 nt)
9082-9128 <--> 11973-12019 51% (47 nt)
9129-9153 <--> 12021-12045 64% (25 nt)
9160-9202 <--> 12046-12088 58% (43 nt)
9203-9292 <--> 12090-12179 73% (90 nt)
9293-9313 <--> 12181-12201 62% (21 nt)
```

```
Seq 1 = ">HS_IL-2Ra_9.3KB genomic fragment from ENSG00000134460 5'->3'"
Seq 2 = ">MmIL-2Ra_12.5kb genomic fragment from phage Ch4C19 5'->3'"

         Local Alignment Number 6
         Similarity Score:  24894
         Match Percentage:  71 %
         Number of Matches:  425
         Number of Mismatches:  142
         Total Length of Gaps:  29
         Begins at (8724,11629) and Ends at (9313,12201)
    0     .    :    .    :    .    :    .    :    .    :
 8724 GAGGACTCAGCTTATGAAGTGCTGGGTGAGACCACTGCCAAGAAGTGCTT
      ||||||||||:|||::||:: ||::|||||||||||||||||||||||||
11629 GAGGACTCAGTTTACAAAACCCTAAGTGAGACCACTGCCAAGAAGTGCTT
   50     .    :    .    :  kB site   :    .  SRE    .SP1/
 8774 GCTCACCCTACCTTCAACGGCAGGGGAATCTCCCTCTCCTTTTATGGGCG
      ||||||||: ||| | :||||||||-||||:||||:||||| ||::|||:
11679 GCTCACCCCTCCTGCCGCGGCAGGG AATCCCCCTTTCCTTGTACAGGCA
  100 GC-box   :    .    :    .    :    .    :    .    :
 8824 TAGCTGAAGAAAGGATTCATAAATGAAGTTCAATCCTTCTCATCAACCCC
       |:|  ||:||||||:||||||:|||||::::||||||||||:||| | |
11728 AAACACAAAAAAGGACTCATAAGTGAAGCCTGATCCTTCTCACCAAACAC
  150     .    :    .    :    .    :    .    :    .    :
 8874 AGCCCACACCTCC AGCAATTGAACTTGAAAAAAAAAACCTGGTTTGAAA
       ||||||||||||-||:||||||||||||||||||||||-||||||||||
11778 TGCCCACACCTCCTAGTAATTGAACTTGAAAAAAAAAAC TGGTTTGAAA
  200     .    : HMGI(Y)        .HMGI(Y)   .    :    .EBS/:
 8923 AATTACCGCAAACTATATTGTCATCaaaaaaaaaaaaaaaaaaaaaCACT
      |||||||||||||:||||||||||-------|||||||||||||||||||
11827 AATTACCGCAAACCATATTGTCAT       AAAAAAAAAAAAAAACACT
  250 Elf-1 HMGI(Y) .    :    .    :    .    :    .    :
 8973 TCCTATATTTGAGATGAGAGAAGAGAGTGCTAGG  CAGTTTCCTGGCTG
      |||||||--|||||| | |||| ||||---||||--||: |||||-||||
11870 TCCTATA  TGAGATCACAGAACAGAG   TAGGCACAAGTTCCT GCTG
  300     .    :    .TATA box .    :    .    :    .+1  :
 9021 AACACGCCAGCCCAATACTTAAAGAGAGCAACTCCTGACTCCGATAGAGA
      |:|| ::|||||:|||:|||||| |||:|||||||||:|| : || || |
11914 AGCAGATCAGCCTAATGCTTAAATAGAACAACTCCTGGCTGTCATTGACA
  350     .    :    .    :    .    :    .    :    .    :
 9071 CTGGATGGACCCACAAGGGTGACAGCCCAGGCGGACCGATCTTCCCATCC
      :||  |::|-- :| |:|:|||||| |:::| || |:|| |::::  ||:
11964 TTGTCTAAA  AGCCAAGATGACAGACTGAGAGGCCTGAGCCCTTGTTCT
  400     .    :    .    :    .    :    .    :    .    :
 9121 CACATCCT CCGGCGCGATGCCAAAAAGAGGCTGACGGCAACTGGGCCTT
       :|||:||-||:| : ||||| : ||||:||:||------||: ::: | 
12012 GGCATTCTCCCAGGAAGATGCAGTAAAGGGGTTG      ACCCAATATA
  450     .    :    .    :    .    :    .    :    .    :
 9170 CTGCAGAGAAAGACCTCCGCTTCACTGCCCCGG CTGGTCCCAAGGGTCA
      ||||||||||   | |||: ||| || |:||: -|||:||||| | |:||
12056 CTGCAGAGAATTTCATCCAGTTCCCTCCTCCATCCTGATCCCATGTGCCA
  500     .    :    .    :    .    :    .    :    .    :
 9219 GGAAGATGGATTCATACCTGCTGATGTGGGGACTGCTCACGTTCATCATG
      |||||||||| :||::|:||||||||| |||::| ||| |:|| |:|||:
12106 GGAAGATGGAGCCACGCTTGCTGATGTTGGGGTTTCTCTCATTAACCATA
  550     .    :    .    :    .    :    .    :    .
 9269 GTGCCTGGCTGCCAGGCAGGTAAG GGCCTGTGGGTGCCCCCGGAA
      ||:||::|:||:|:||||||||||-|:|| | :: |||||: ||||
12156 GTACCCAGTTGTCGGGCAGGTAAGAGACCAGGAACTGCCCTGGGAA
```

PRRI
[-276,-244]

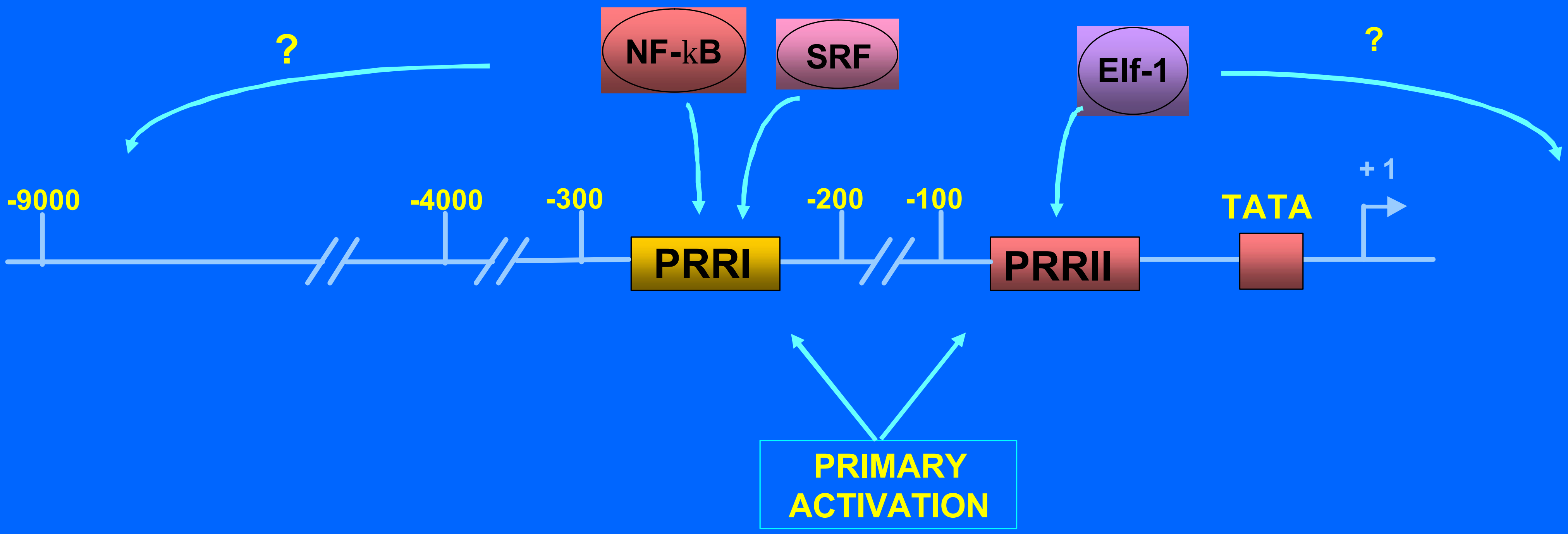
PRRII
[-137,-64]

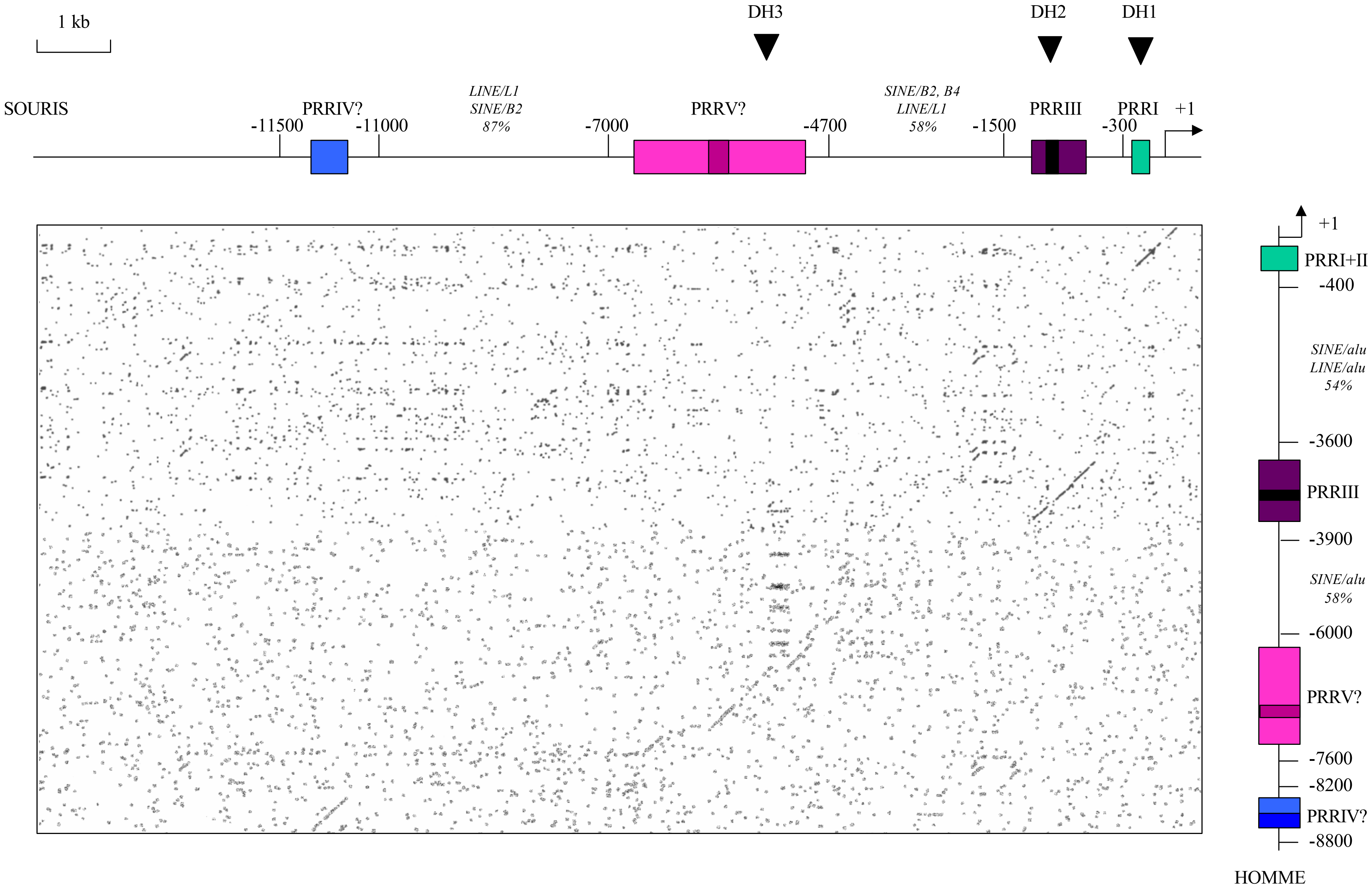
HOWEVER NEITHER PRRI NOR PRRII, ALONE OR IN ASSOCIATION, ARE SUFFICIENT

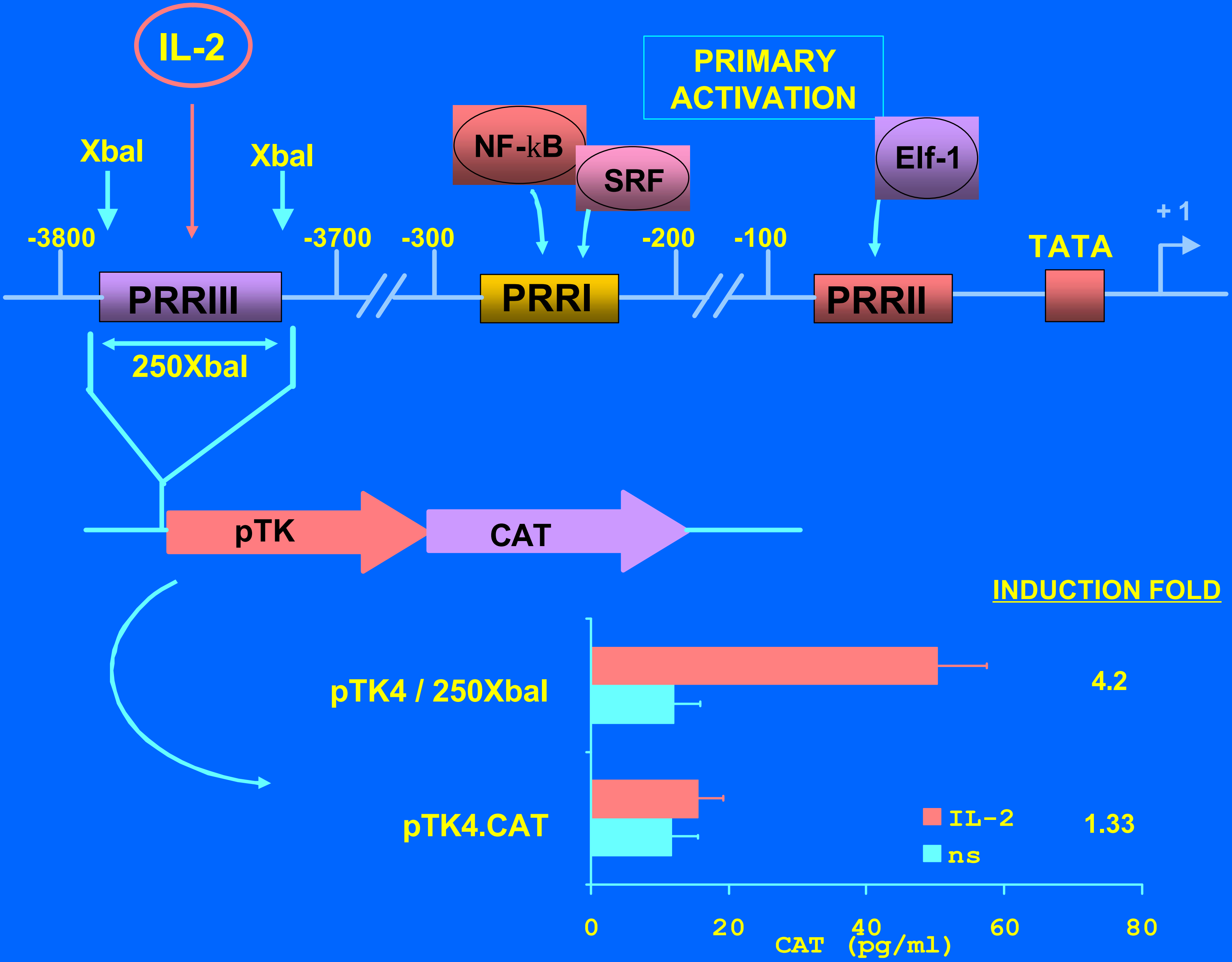
TO ELICIT AN EFFICIENT TRANSCRIPTION

IN GENE REPORTER ASSAYS USING VARIOUS T CELL MODELS

SEARCH FOR IL-2 AND/OR CD28 RESPONSIVE REGULATORY REGIONS WITHIN CD25/IL-2Rα GENE LOCUS




1 kb
DH3
DH2
DH1
SOURIS
PRRIV?
LINE/L1
SINE/B2
87%
PRRV?
SINE/B2, B4
LINE/L1
58%
PRRIII
PRRI
+1
-11500
-11000
-7000
-4700
-1500
-300
+1
PRRI+II
-400
SINE/alu
LINE/alu
54%
-3600
PRRIII
-3900
SINE/alu
58%
-6000
PRRV?
-7600
-8200
PRRIV?
-8800
HOMME

IL-2
PRIMARY
ACTIVATION
XbaI
XbaI
NF-kB
SRF
Elf-1
+ 1
-3800
-3700
-300
-200
-100
TATA
PRRIII
PRRI
PRRII
250XbaI
pTK
CAT
INDUCTION FOLD
pTK4 / 250XbaI
4.2
pTK4.CAT
1.33
IL-2
ns
0
20
40
60
80
CAT (pg/ml)

PRRIII

Homo Sapiens

GASd/EBSd   GASp/GATA   EBSp

```
- 3772                                                            - 3718
        TTTCTTCTAGGAAGTACCAAACATTTCTGATAATAGAATTGAGCAATTTCCTGAT
        IIIIIIII--IIIIIIIII-IIIIIIIIIIIII-III-IIIIIIII-IIIIIIII
        TTTCTTCTGAGAAGTACCAGACATTTCTGATAAGAGAGTTGAGCAACTTCCTGAT
- 1369                                                            - 1315
```

Site I   Site II   Site III

Mus Musculus

IL-2rE

In Vivo Footprinting

- 1) Methylation of guanines (major groove) and to a lesser extent of adenines (minor groove) by DMS on living cells

  The level of methylation is affected by protein binding to DNA

- 2) Genomic DNA extraction

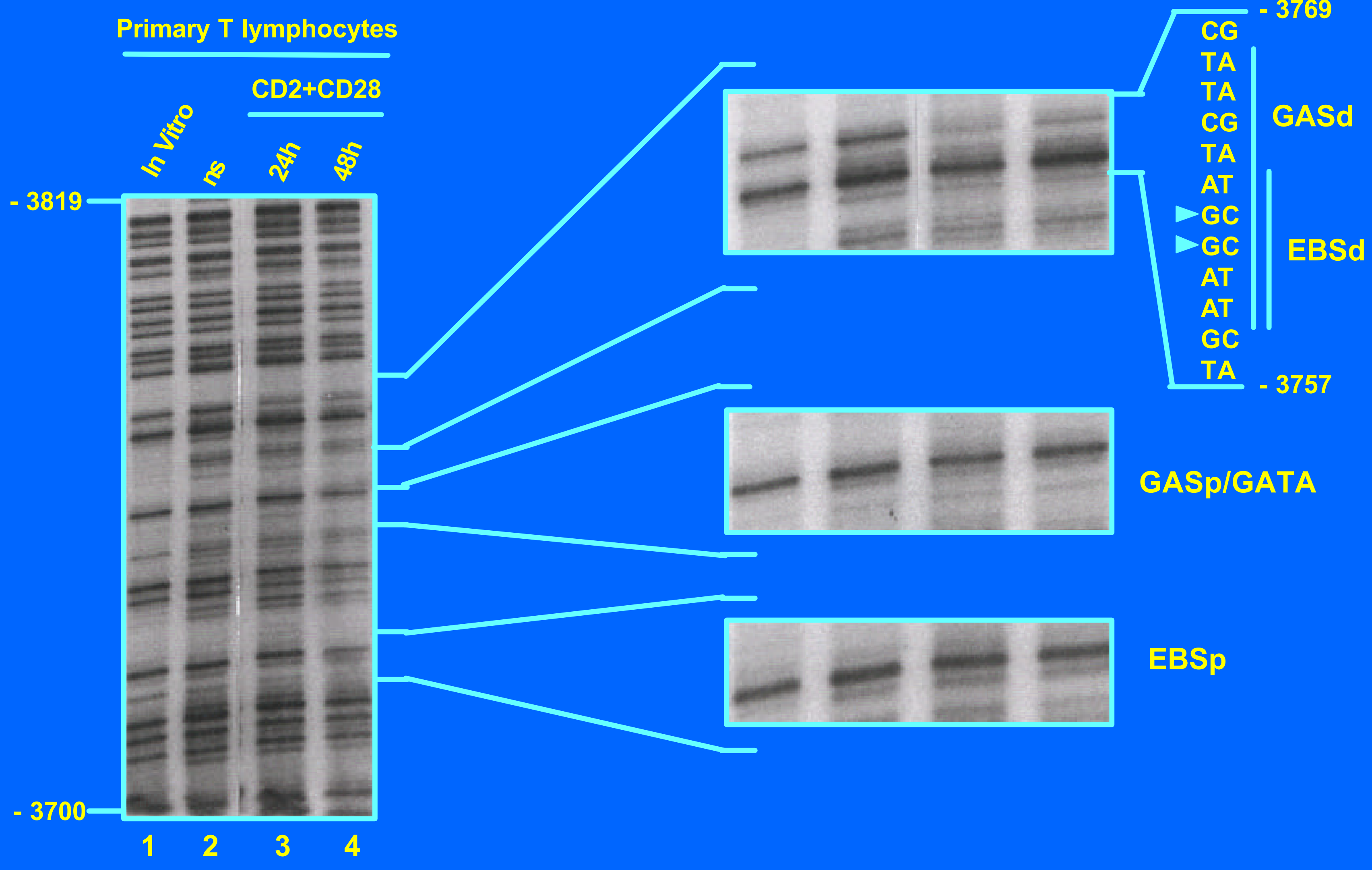
- 3) Cleavage of methylated residues by piperidine

- 4) LMP-PCR amplification of the region to be analyzed

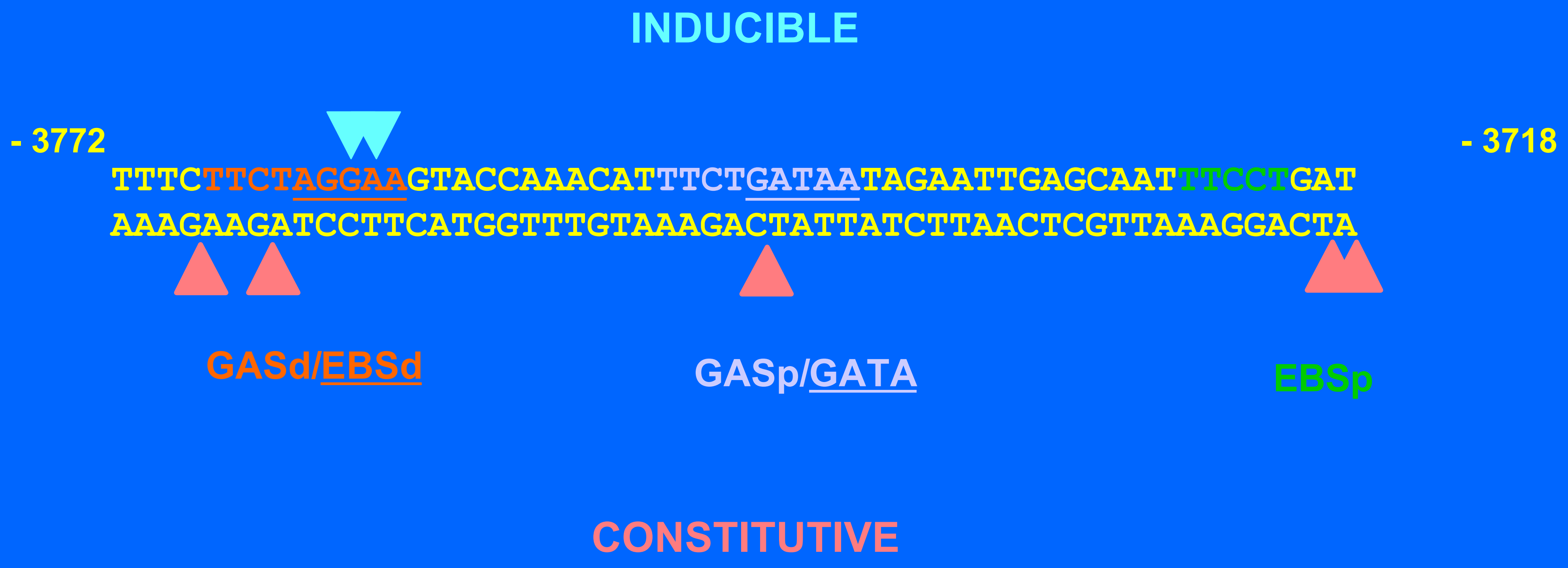
  Last amplification cycles performed with a ^{32}P-labeled primer

- 5) Analysis of the PCR products on sequencing gel

In vivo modification of the GASd/EBSd motif occupancy in response to CD2+CD28 costimulation in purified human primary T cells



The GASd/EBSd motif is the only putative regulatory element within PRRIII modified *in vivo* in response to an IL-2-dependent induction in human T lymphocytes



INDUCIBLE

- 3772

- 3718

TTTCTTCTAGGAAGTACCAAACATTTCTGATAATAGAATTGAGCAATTTCCTGAT

AAAGAAGATCCTTCATGGTTTGTAAAGACTATTATCTTAACTCGTTAAAGGACTA

GASd/EBSd

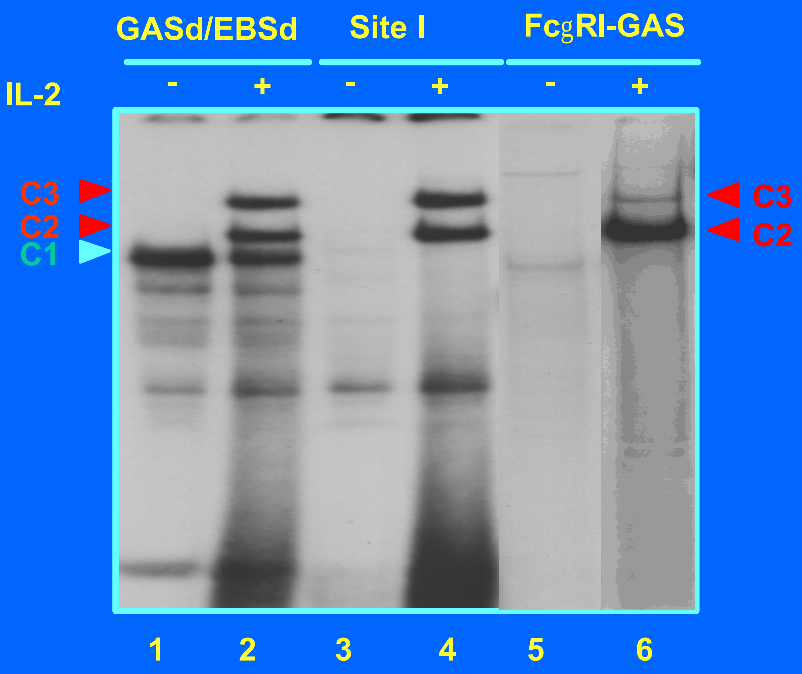
GASp/GATA

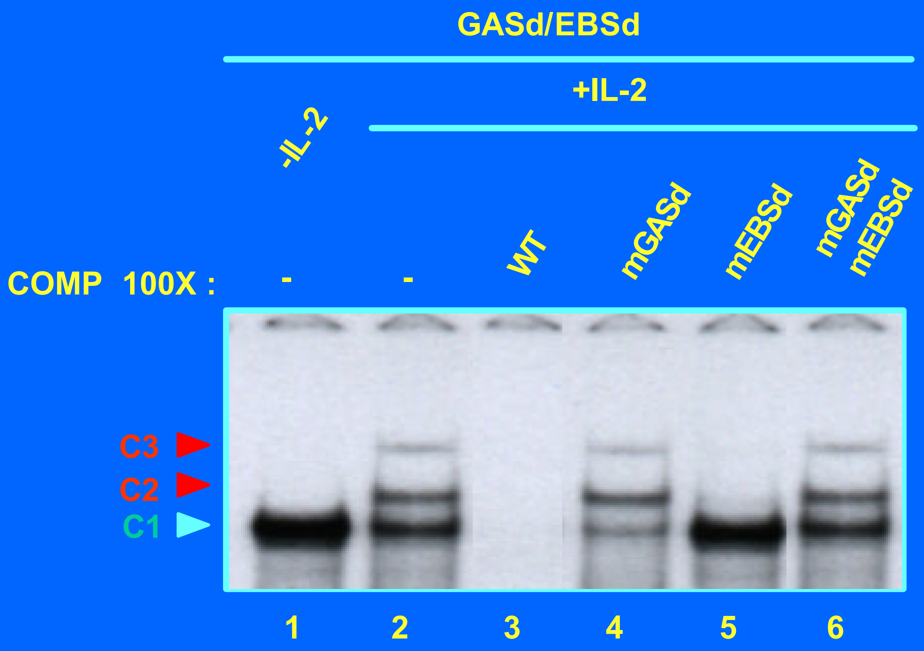
EBSp

CONSTITUTIVE

Lecine, P., Algarte, M., Rameil, P., Beadling, C., Bucher, P., Nabholz, M. and Imbert, J. Elf-1 and Stat5 bind to critical element in a new enhancer of the human interleukin-2 receptor alpha gene. Mol.Cell.Biol. 16: 6829-6840; 1996.

Inducibles complex C2 and C3 are GAS-specific
Constitutive complex C1 is EBS-specific





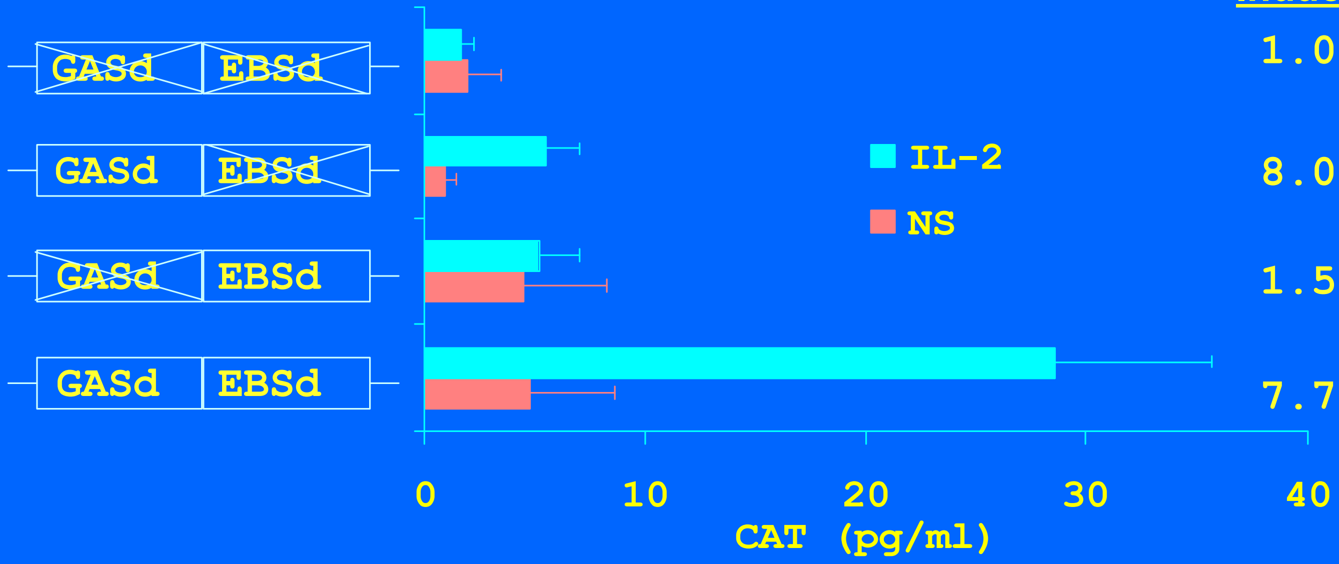
CD25/IL-2Ra GASd/EBSd EMSA probe:

GASd

```
TTTCTTCTAGGAAGTACC
AAAGAAGATCCTTCATGG
```

EBSd

Disruption of the GASd motif by site-directed mutagenesis fully abolish IL-2 response



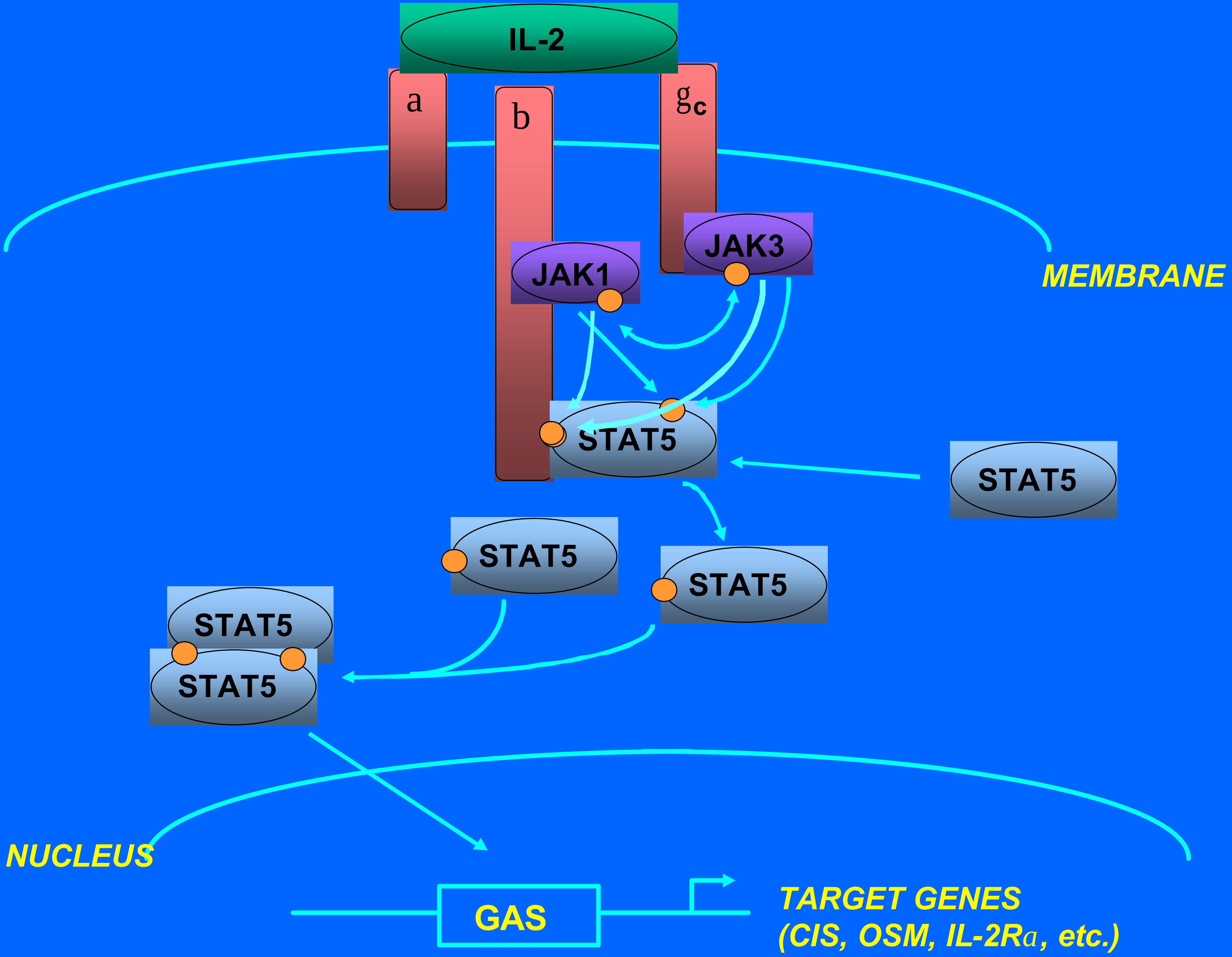
Induction Fold

1.0 +/- 0.1

8.0 +/- 3.7

1.5 +/- 0.1

7.7 +/- 2.4



IL-2

a

b

g_c

MEMBRANE

JAK1

JAK3

STAT5

STAT5

STAT5

STAT5

STAT5

STAT5

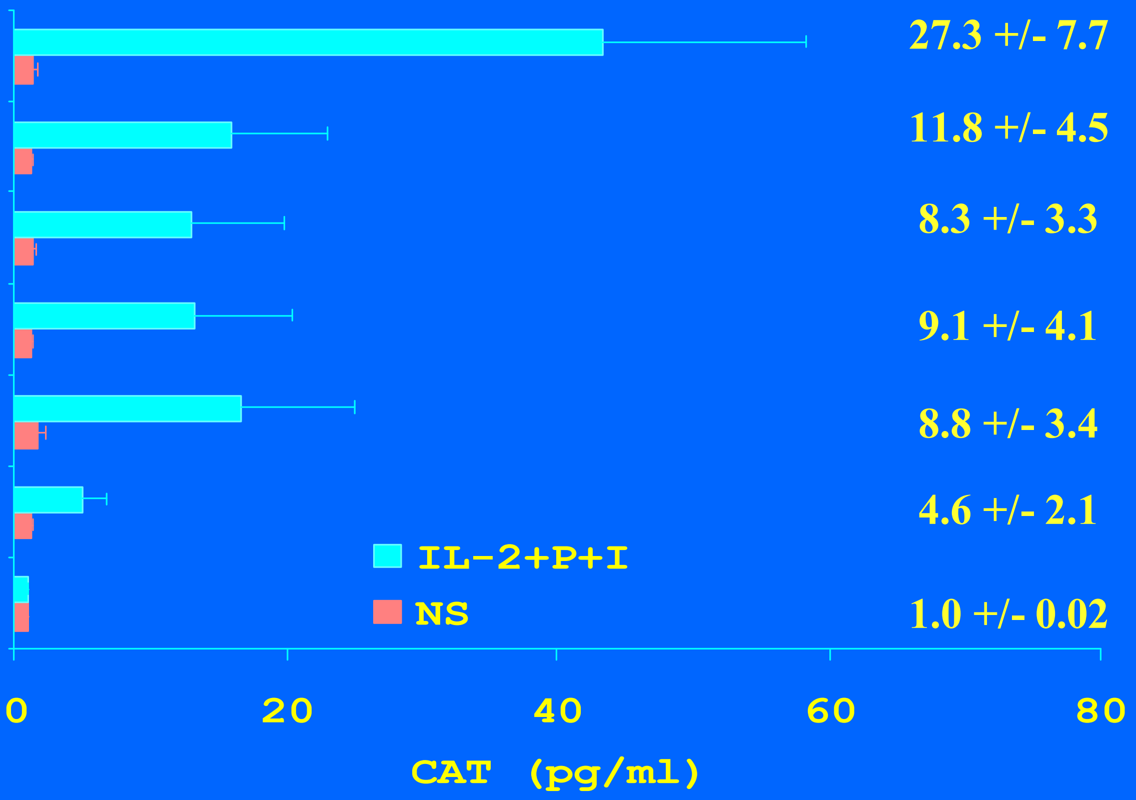
NUCLEUS

GAS

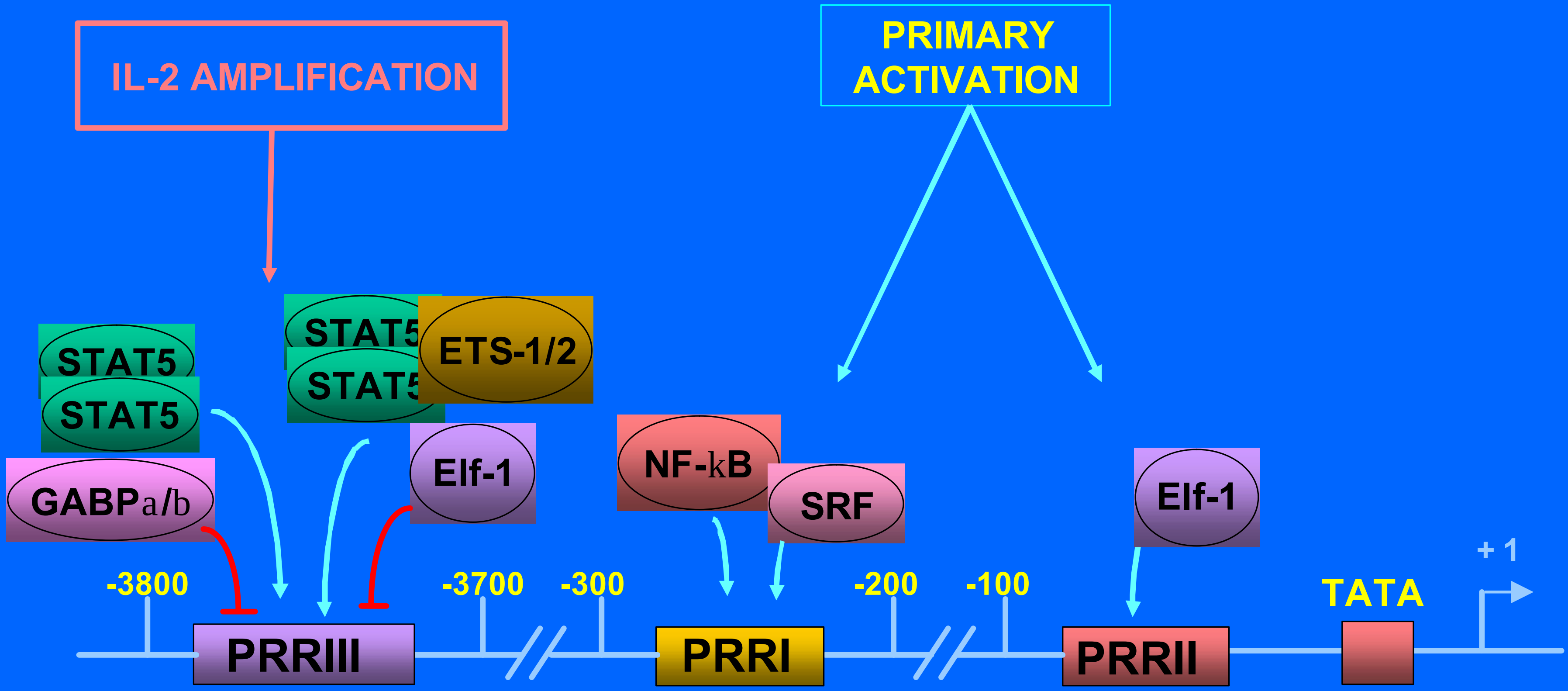
TARGET GENES
(CIS, OSM, IL-2Ra, etc.)

Stat5b, Ets-1 and Ets-2 cooperate functionally in response to IL-2 +PMA+Ionomycin

REPORTER	EFFECTOR Stat5b	EFFECTOR Ets-1	EFFECTOR Ets-2	INDUCTION FOLD
GASd EBSd	+	+	+	27.3 +/- 7.7
GASd EBSd	-	+	+	11.8 +/- 4.5
GASd EBSd	-	-	+	8.3 +/- 3.3
GASd EBSd	-	+	-	9.1 +/- 4.1
GASd EBSd	+	-	-	8.8 +/- 3.4
GASd EBSd	-	-	-	4.6 +/- 2.1
pTK4.CAT	-	-	-	1.0 +/- 0.02



Rameil, P., Lecine, P., Ghysdael, J., Gouilleux, F., Kahn-Perles, B. and Imbert, J. IL-2 and long-term T cell activation induce physical and functional interactions between STAT5 and ETS transcription factors in human T cells. Oncogene 19: 2086-2097; 2000.



IL-2 AMPLIFICATION

PRIMARY ACTIVATION

STAT5

STAT5

GABPa/b

STAT5

STAT5

ETS-1/2

Elf-1

NF-kB

SRF

Elf-1

-3800

-3700

-300

-200

-100

TATA

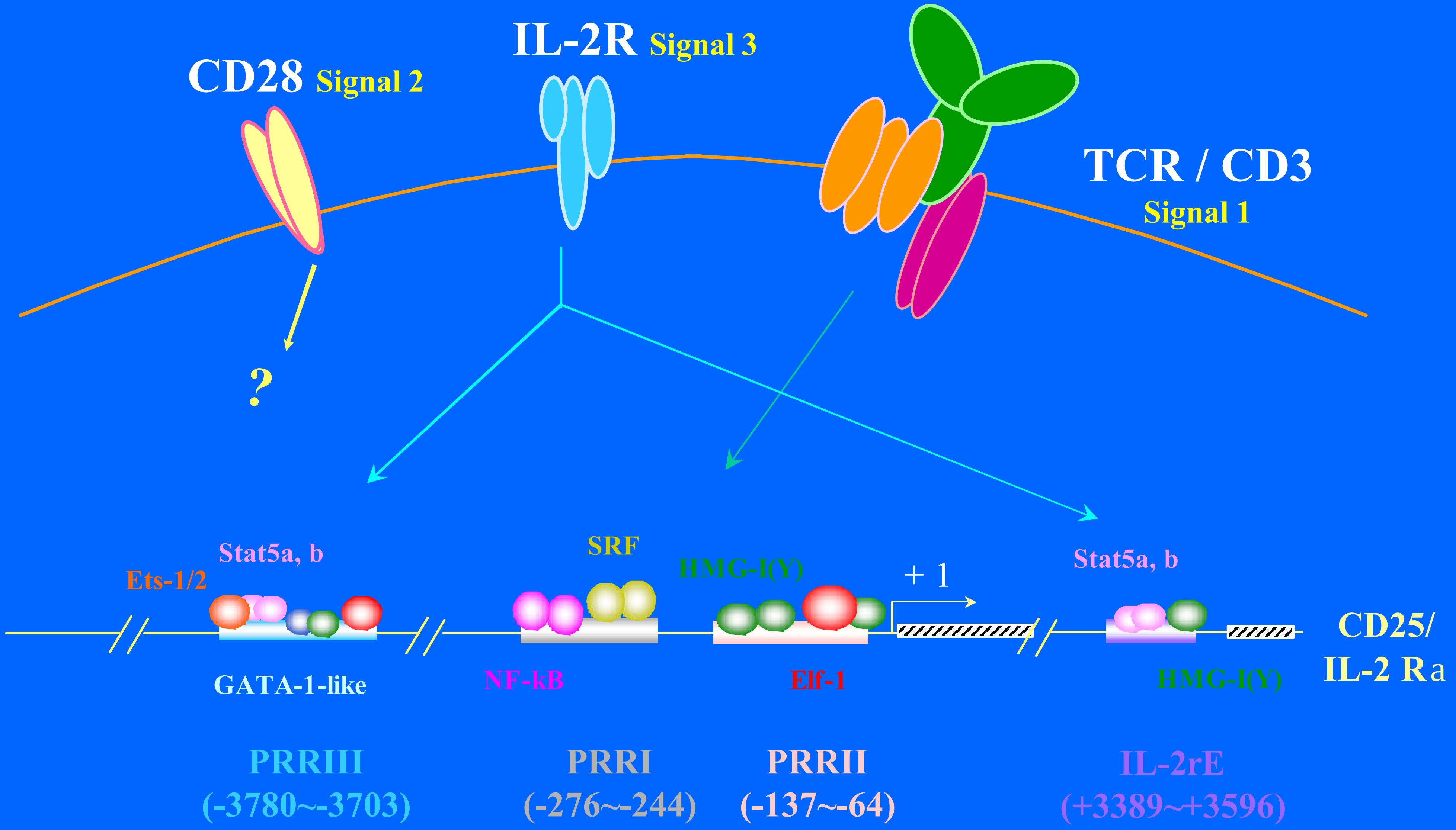
+ 1

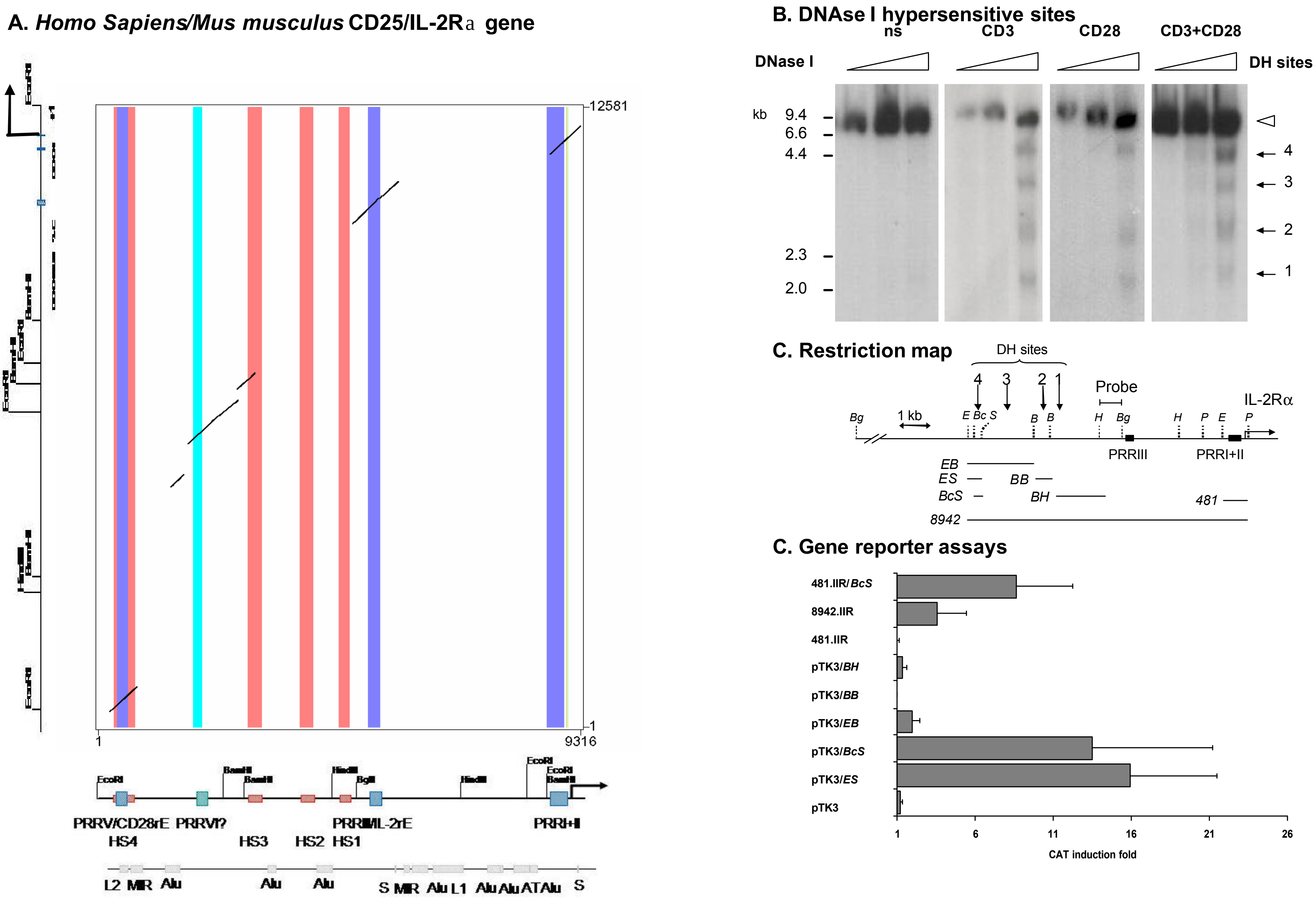
PRRIII

PRRI

PRRII

Regulatory elements and their cognate transcription factors in the human CD25/IL-2Ra gene





A. *Homo Sapiens/Mus musculus* CD25/IL-2Rα gene

B. DNAse I hypersensitive sites

C. Restriction map

C. Gene reporter assays

A Well-Conserved but not Functional Candidate for a CD28 Responsive Enhancer

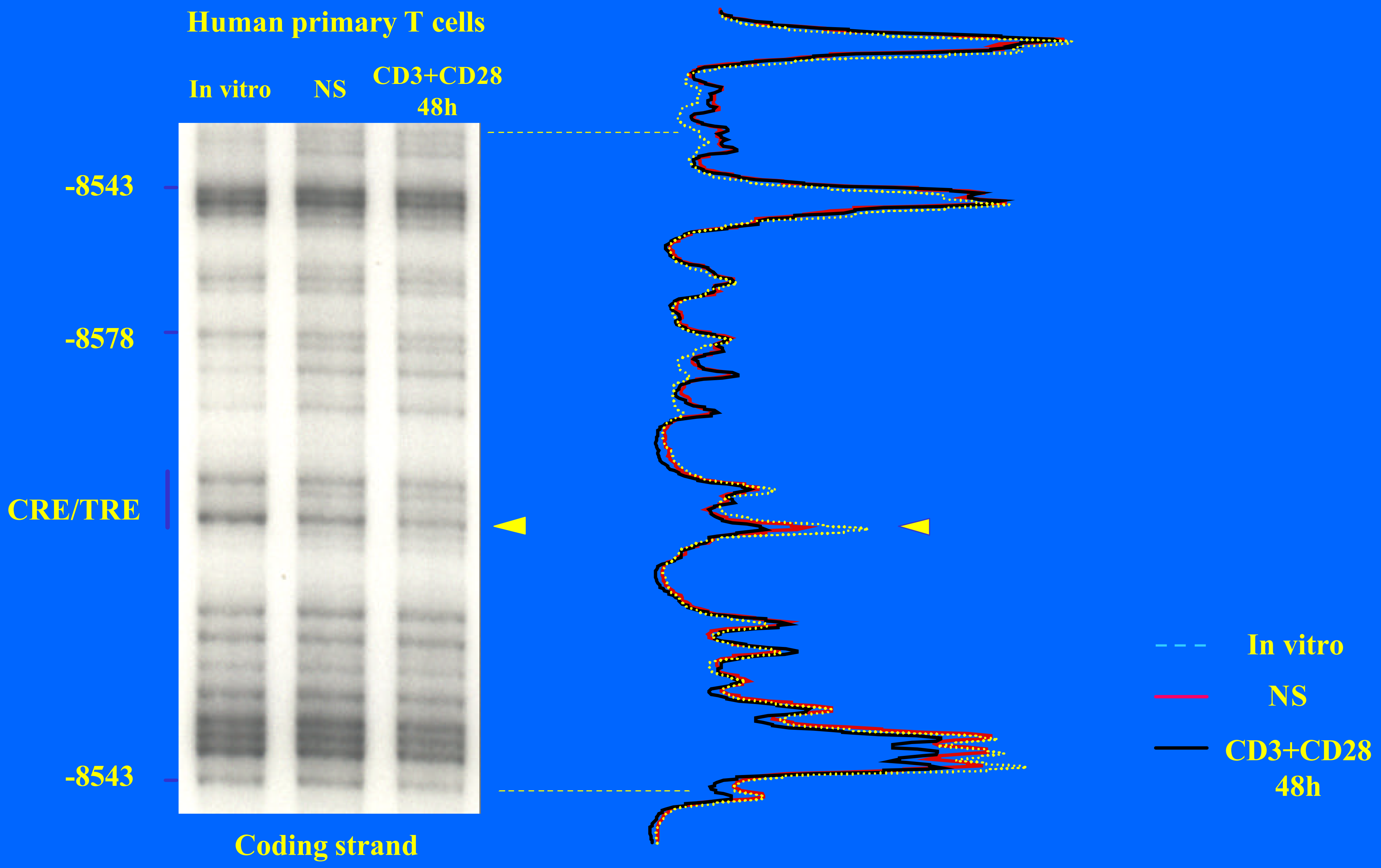
Comparison of two nucleotidic sequences:
Sequence 1 : cons4HS
Sequence 2 : cons4mus

```
 resetting to DNA matrix
 LALIGN finds the best local alignments between two sequences
 version 2.0u4 Feb. 1996
Please cite:
 X. Huang and W. Miller (1991) Adv. Appl. Math. 12:373-381

 Comparison of:
(A) cons4HS
(B) cons4mus
 using matrix file: DNA, gap penalties: -16/-4

  79.6% identity in 201 nt overlap; score:  612

                       NF-ATp
               10        20        30        40        50        60
cons4H CATAGTGGATTTTGGTTTTCCACGGGACCCCTGTGCCCTTGTCTAGTAGAATCTGGTGGA
       ::::::::::: ::::::::::: ::::::         ::::::::: :: :: :::::
cons4m CATAGTGGATTCTGGTTTTCCACAGGACCC---------TGTCTAGTAAAACCTAGTGGA
               10        20        30                 40        50
                    NF-kB/CD28RC  CREB         Ets
               70        80        90       100       110       120
cons4H AATTACAAACTGCAGAAATTCAACTCAGTGCCGCAATAACAGGATGCACCTGTAGATTTC
       :::::::: ::: ::::::::: :   ::::: :::::::::::::::::::::::::::
cons4m AATTACAAGCTG-AGAAATTCAGCCTTGTGCCACAATAACAGGATGCACCTGTAGATTTC
              60         70        80        90       100       110
       GAS
              130       140       150       160       170       180
cons4H GTAGAATTAGCAGCAGCATTCTTTCAATACCAGTTTGAGAGAAATAACCCTGTTTGCATA
         ::::::::: :: :  :::: : ::  ::: ::::::::::: :: ::::::::  : 
cons4m ACAGAATTAGCTGCTGTCTTCTCTTAACGCCAATTTGAGAGAAAGAAGCCTGTTTGTCTG
              120       130       140       150       160       170

              190       200
cons4H GTGCCAACTGGGGCAGAATCT
        ::::::   :::::::::::
cons4m CTGCCAAACAGGGCAGAATCT
              180       190
```

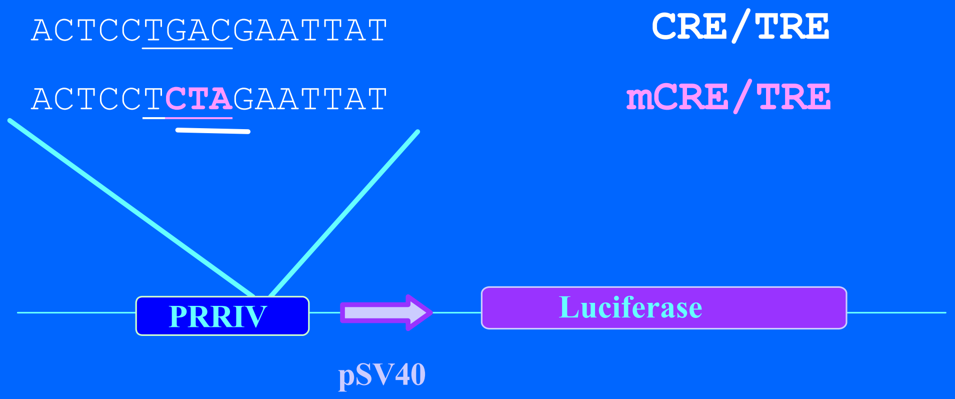
The Functional PRRV/CD28rE within Human CD25/IL-2Ra Gene

```
         Bcl-I    .         .         .         .         .
-8688    TGATCAGCCGTGTCTCCAGAGAGCTACAAGGCAGTTTTCAATTGGTAAAT
         ACTAGTCGGCACAGAGGTCTCTCGATGTTCCGTCAAAAGTTAACCATTTA
                                            NFAT

                  .         .         .         .         .
-8638    GCCCTGAGAGTGATGGGCTTGTGGCATGTGTAAGGGTTAGACAGACCTGG
         CGGGACTCTCACTACCCGAACACCGTACACATTCCCAATCTGTCTGGACC
                  .TRE      .CRE/TRE  .         .         .
-8588    GACTAGACATGACACCACTCCTGACGAATTATGTGAGTGTGGGTGTTTCA
         CTGATCTGTACTGTGGTGAGGACTGCTTAATACACTCACACCCACAAAGT
                   TCF 11      TCF 11
                                             NFAT         
                  .         .         .      NFAT         .
-8538    CAACCACAATGAGATGCAATGCCTGCACTTGTAACATGGAAATAGTGATG
         GTTGGTGTTACTCTACGTTACGGACGTGAACATTGTACCTTTATCACTAC
                  TCF 11

         SphI
-8488    GCATGC
         CGTACG
```

In vivo LM-PCR characterized the regulatory elements in PRRIV/IL-2Ra

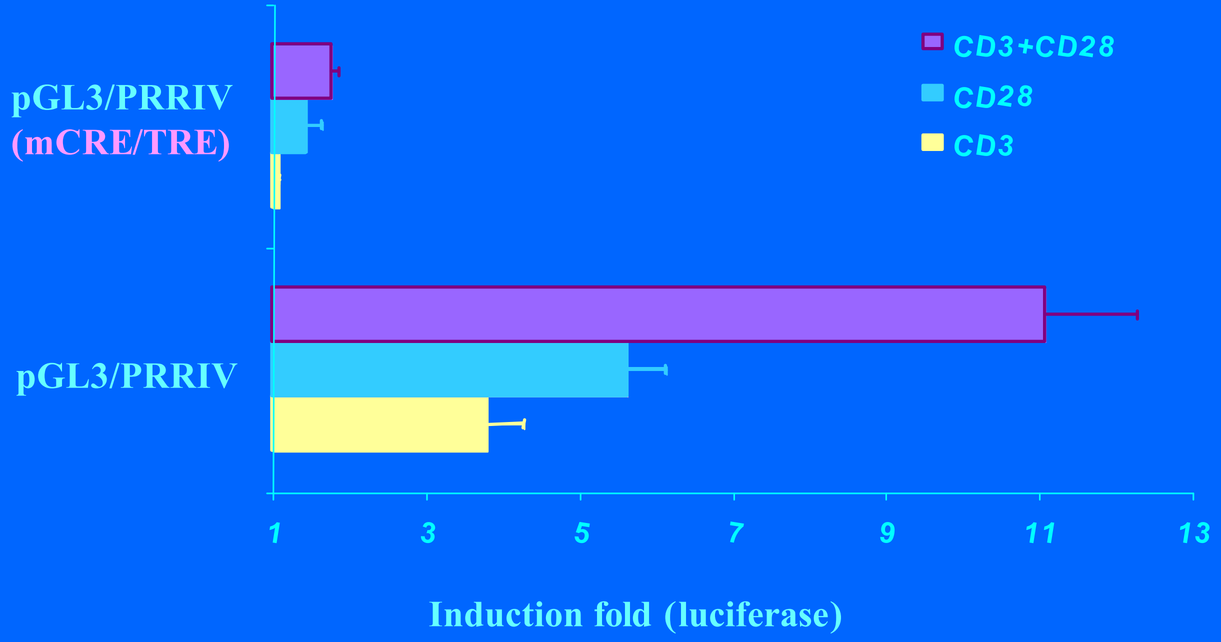


The CRE/TRE within PRRIV is essential for the response of PRRIV to TCR-CD3 and CD28 signals



ACTCCTGACGAATTAT CRE/TRE

ACTCCTCTAGAATTAT mCRE/TRE



The crucial CRE/TRE motif within CD25/IL-2Ra PRRV/CD28rE is contained in a SINE/MIR repeat and is not conserved between *Homo Sapiens* and *Mus Musculus*

```
Local identities (aligment n°1)

211-245 <--> 312-346 74% (35 nt)
261-312 <--> 347-398 62% (52 nt)
313-322 <--> 400-409 70% (10 nt)
323-371 <--> 414-462 61% (49 nt)
374-388 <--> 463-477 60% (15 nt)
389-444 <--> 480-535 75% (56 nt)
447-482 <--> 536-571 53% (36 nt)
491-542 <--> 572-623 67% (52 nt)
544-602 <--> 624-682 61% (59 nt)
603-664 <--> 684-745 66% (62 nt)
666-686 <--> 746-766 57% (21 nt)
690-709 <--> 767-786 85% (20 nt)
712-744 <--> 787-819 58% (33 nt)
```

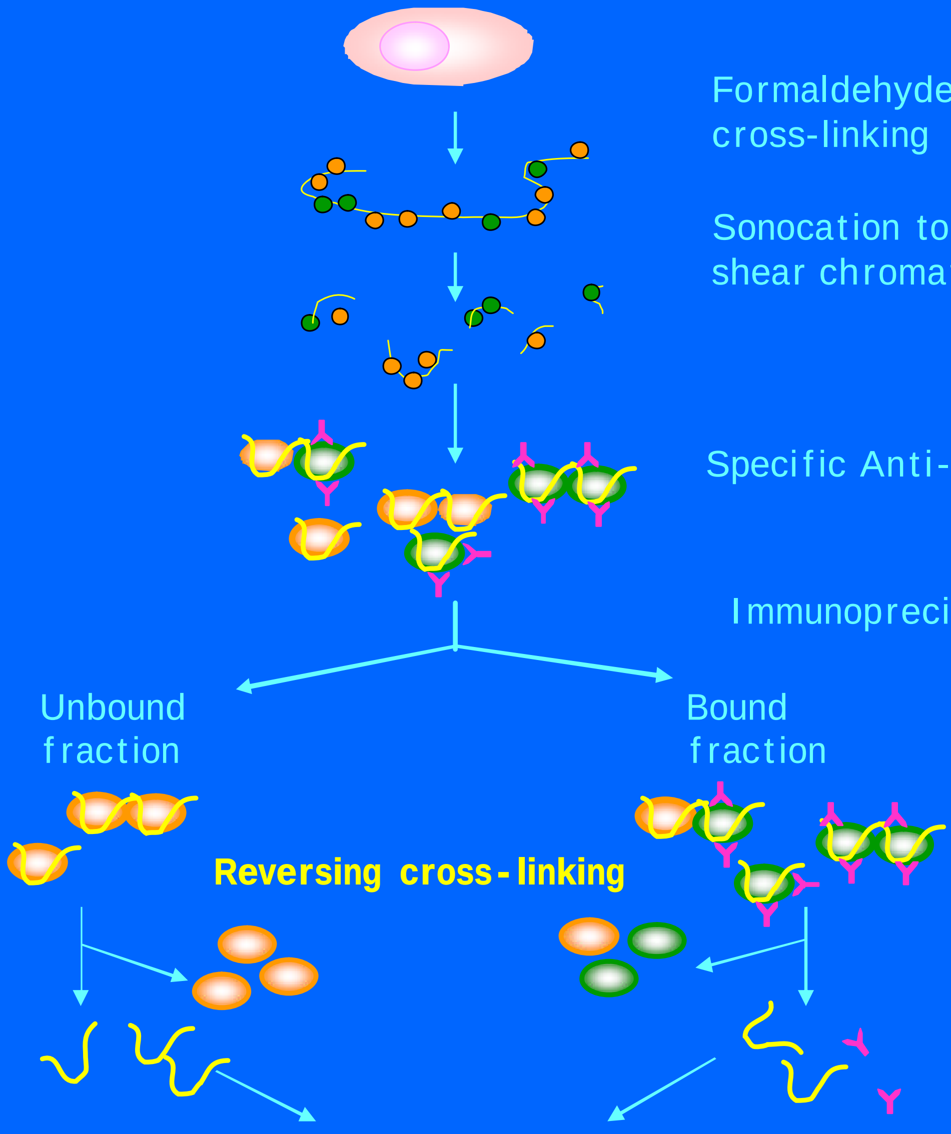
```
Seq 1 = ">HS_IL-2Ra_9.3KB genomic fragment from ENSG00000134460 5'->3'"
Seq 2 = ">MmIL-2Ra_12.5kb genomic fragment from phage Ch4Cl9 5'->3'"

        Local Alignment Number 1
        Similarity Score:  12780
        Match Percentage:  59 %
        Number of Matches:  325
        Number of Mismatches:  175
        Total Length of Gaps:  42
        Begins at (211,312) and Ends at (744,819)

   0    .    :    .    :    .    :    .    :    .    :
 211 CTCTAAAAAGGTTCTGCATACAGtcattcattcaatgcttaacgactgag
     ||:| ||: |||||| ||:| |:||||| ||||||---------------
 312 CTTTTAAGTGGTTCTCCACAGAATCATTAATTCAA

  50    .    :    .    :    .    :    .    :    .    :
 261 cattaattccatgctaagtactgaactcagcactaggaataagaaggcga
     ||   ||: ::|||||:|| ||: |||||| |||::|| |||:| ||::|
 347 CAGGTATCATGTGCTAGGTCCTATACTCAGAACTGAGACTAAAATGGTAA

 100    .    :    .    :    .    :    .    :    .    :
 311 cc tagaggcata    tcctctctctaaagatgcatagagcctcattgg
     | -||||:| | |----|:||:|:||||||:|::||:||||| : ::|||
 397 CAATAGAAGAAGACAATTTCTTTTTCTAAAAACACACAGAGCACAGCTGG

 150   BclI  :    .    :    .    :    .    :    .    :
 356 aatgatcagccgtgtctccagagagctacaagg  cagTTTTCAATTGGT
     |||  |::| |:||| --||||  | |: |:||--||||||| |:|||||
 447 AATTCTTGGGCATGTA  CAGATTGATGGAGGGTACAGTTTTAAGTTGGT

 200    .    :    .    :    .    :    .    :    .    :
 404 AAATGCCCTGAGAGTGATGGGCTTGTGGCATGTGTAAgggttagacagac
     |||||:|:||||||: |||:  | ||:|| ||| |:|||||--||::: :
 495 AAATGTCTTGAGAGCTATGACATGGTAGCCTGTTTGAGGGT  GATGATT

 250    .    :    .    :  CRE/TRE:    .    :    .    :
 454 ctgggacctagacatgacaccactcctgacgaattatgtgagtgtgggtg
      |:|:  |||  | | |||:| ||||||:---------||||| |||||::
 543 GTAGACACTACTCCTCACATCTCTCCTGG        GTGAGAGTGGGCA

 300    .    :    .    :    .    :    .    :    .    :
 504 tttcacaaccacaatgagatgcaatgcctgcacttgtaacatggaaatag
     |||||:||::|||:|:|:||::||||:|| ||:: |:||-|||||||| |
 585 TTTCATAATTACAGTAAAATATAATGTCTTCATCAGCAA ATGGAAATCG

 350    .SphI    .    :    .    :    .    :    .    :
 554 tgatggcatgccggccccgccagattgctgtgagaagtcagcggcagag
     |||:::::||:|:||:  :|  |:||||||||| |||:||||::: | |:-
 634 TGACAATGTGTCAGCTAGACATGGTTGCTGTGTGAAATCAGTAAGATAAC

 400    .    :    .    :    .    :    .    :    .    :
 603 acatgcaacattctcagcacagtgcttgccatgtagtaagggcctagtca
     ||||  || || : ::|||||| ||:|||:|||||:|||: :::|| |||
 684 ACATTAAAAATGTATGGCACAGAGCCTGCTATGTAATAAACATTTATTCA

 450    .    :    .    :    .    :    .    :    .    :
 653 gtgctagTGATTCCTTTCAATATTCCTAAGATGCAGATAAGGGAACAGCC
     :|| |||||:||- |||| :||:|   :| ||||---||||| ||||||:
 734 ATGGTAGTGGTT ATTTCCGTACTAGAGATATGC   TAAGGCAACAGCT

 500    .    :    .    :    .    :    .    :
 703 CAGAGGAGGGGGAGCACTTCCAGAGGGAGGGATGCGGTGAGA
     |||||:|--|:||::||:||:| | :: ||||| :|:|||:|
 780 CAGAGAA  GAGAATACCTCTACATAATGGGATCTGATGAAA
```

HS4 (-8.5Kb)
CD28rE (PRRV)
[-8689,-8484]

BclI-SphI fgt

underlined:
SINE/MIR repeat

Chromatin immunoprecipitation (ChIP)



Formaldehyde cross-linking

Sonocation to shear chromatin

Specific Anti-bodies

Immunoprecipitation

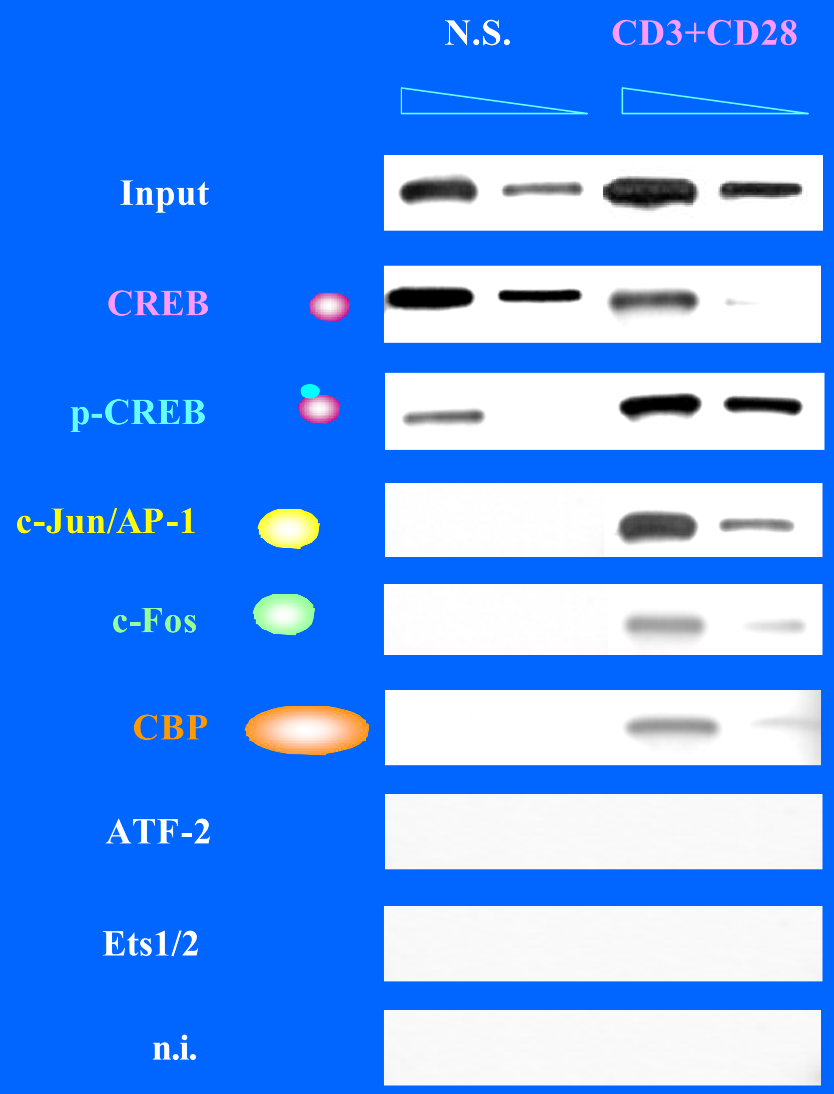
Unbound fraction

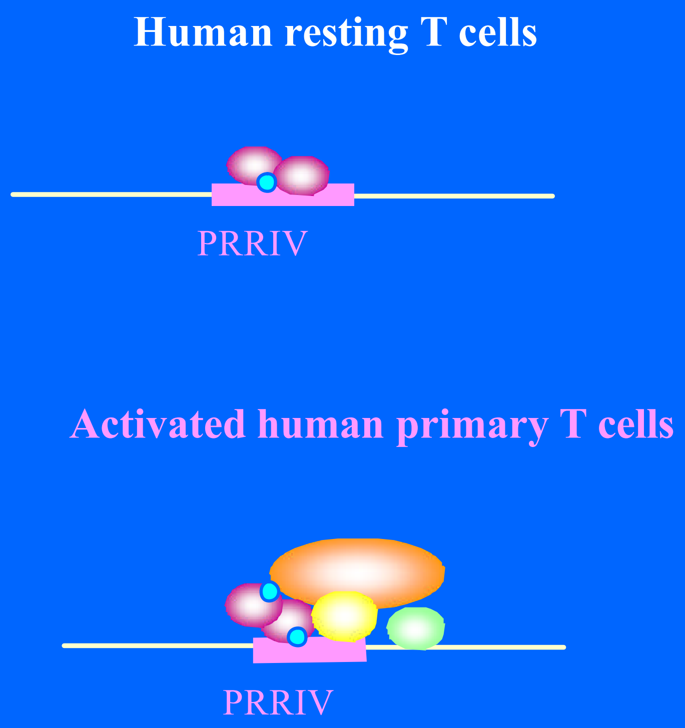
Bound fraction

Reversing cross-linking

Semi-quantitative PCR Assays

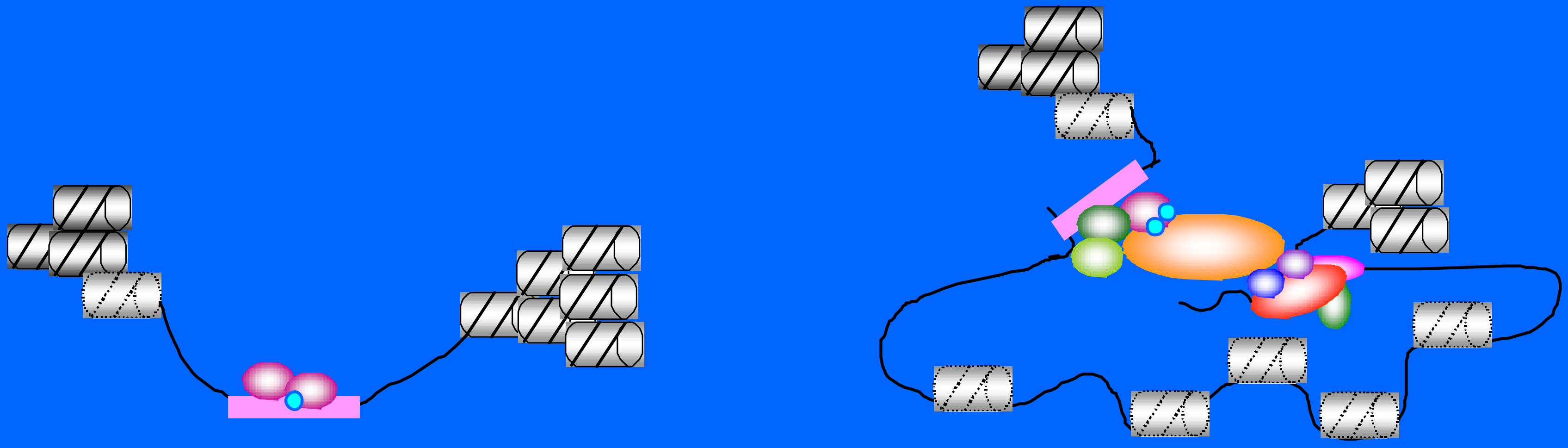
Specific recruitment of CREB, CBP, c-Jun/AP-1, and c-Fos/AP-1 to PRRIV in vivo





Resting human primary T cells

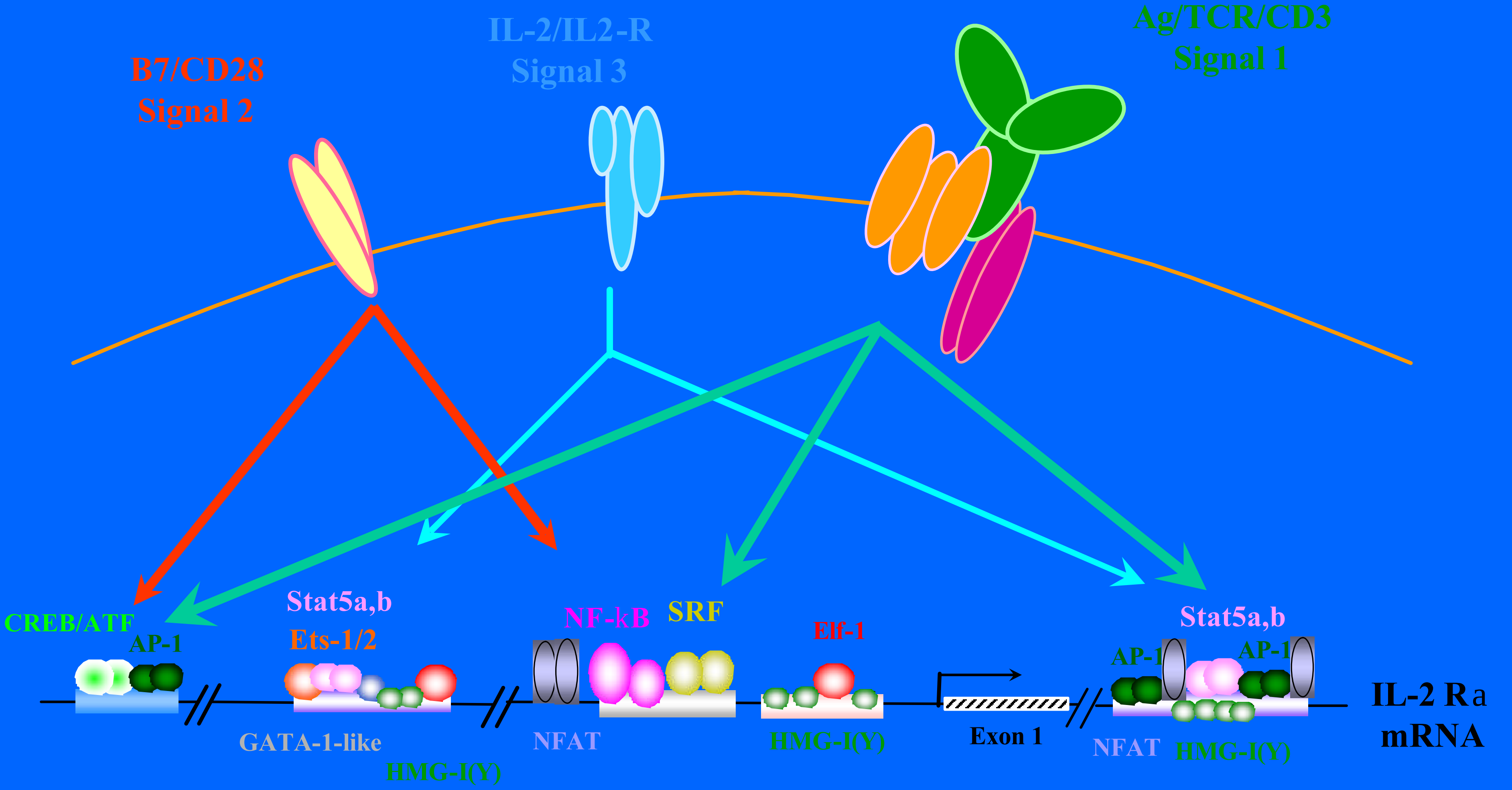
Activated human primary T cells



IL-2Ra gene transcription : Off On

▨	Nucleosome	▬	CRE/TRE	●	c-Jun/AP-1
▒	Remodelled nucleosome	●	CREB	⬭	CBP/p300
●	RNA polymerase II complex	●	Phosphorylated Ser^{133} of CREB	●	c-Fos/AP-1

Yeh, J.H., Lecine, P., Nunes, J.A., Spicuglia, S., Ferrier, P., Olive, D. and Imbert, J. Novel CD28-responsive enhancer activated by CREB/ATF and AP-1 families in the human IL-2Ralpha locus. Mol.Cell.Biol., 21: 4515-4527; 2001.



B7/CD28
Signal 2

IL-2/IL2-R
Signal 3

Ag/TCR/CD3
Signal 1

CREB/ATF
AP-1

Stat5a,b
Ets-1/2
GATA-1-like
HMG-I(Y)

NF-kB
SRF
NFAT

Elf-1
HMG-I(Y)

Exon 1

Stat5a,b
AP-1
AP-1
NFAT
HMG-I(Y)

IL-2 Ra
mRNA

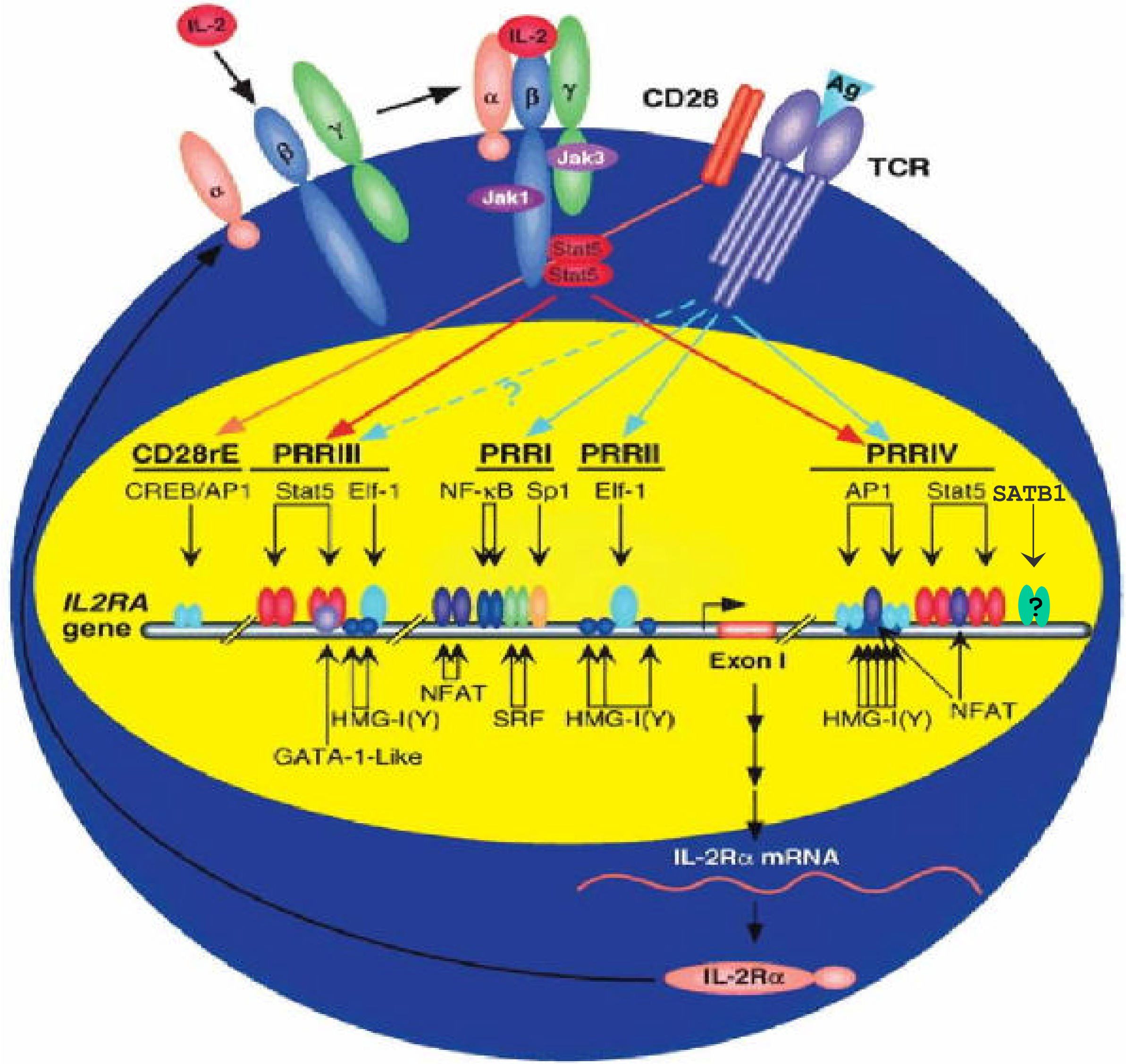
PRRV/CD28rE
(-8689~-8483)

PRRIII/IL-2rE
(-3780~-3703)

PRRI
(-276~-244)

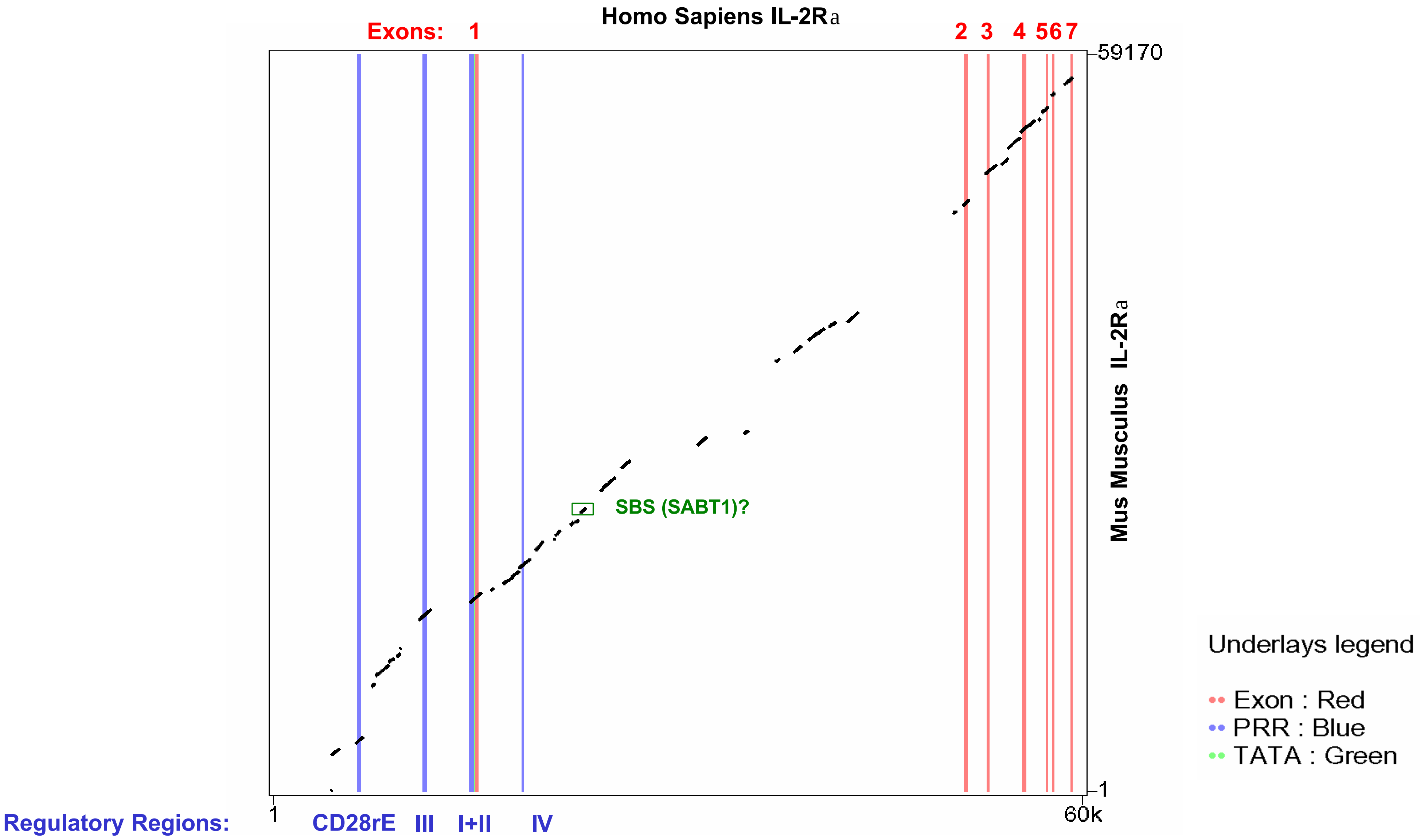
PRRII
(-137~-64)

PRRIV/IL-2rE
(+3389~+3596)



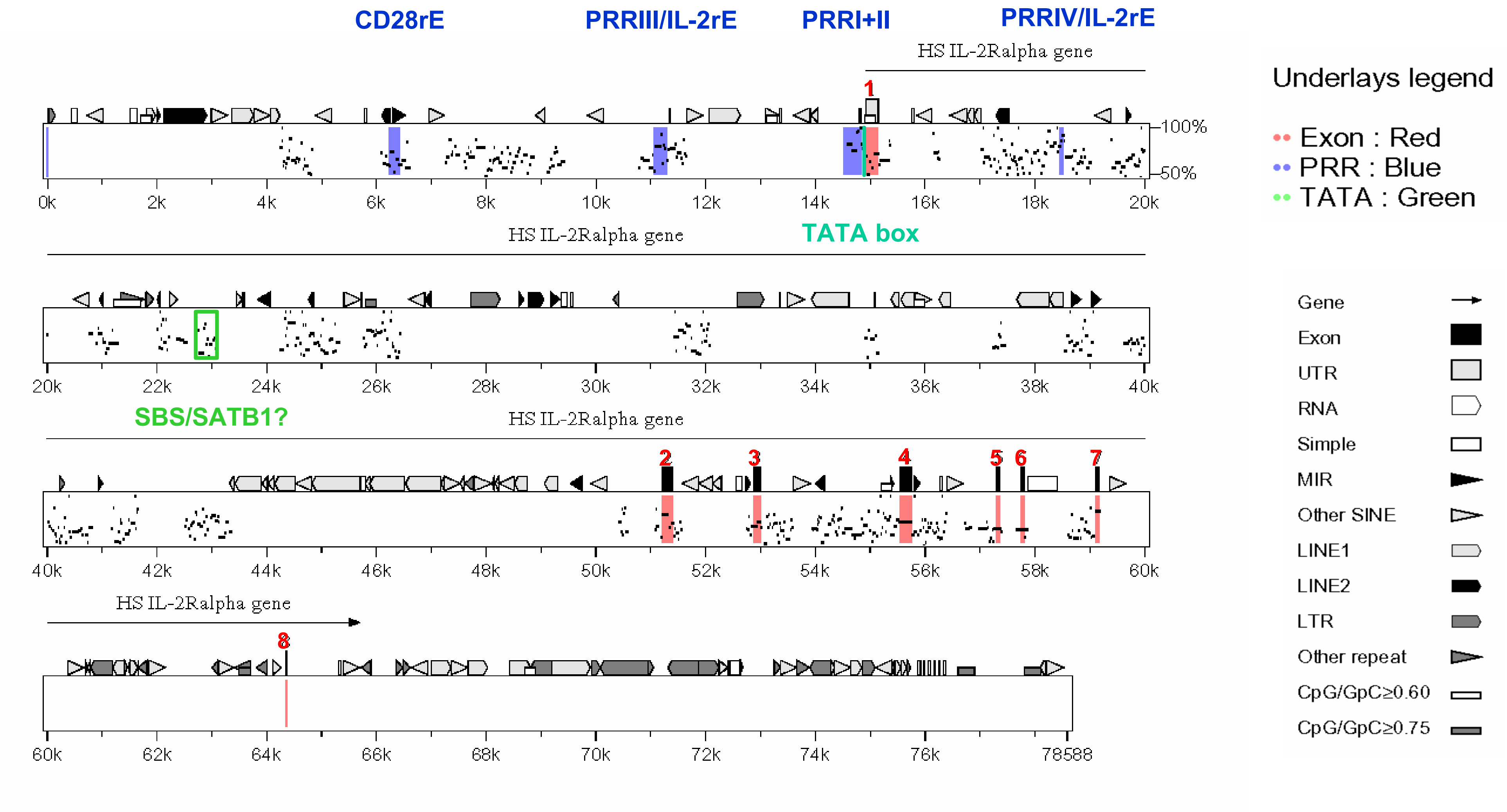
From Kim et al., EMBO J., 21:3051-3059, 2002.

Homo Sapiens/Mus Musculus IL-2Ra locus dotplot comparison



Schwartz, S., Zhang, Z., Frazer, K.A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R., and Miller, W. (2000). PipMaker: A Web Server for Aligning Two Genomic DNA Sequences. Genome Res. 10, 577-586.

Percent Identity Plot (PIP) of Homo Sapiens and Mus Musculus IL-2Rα locus



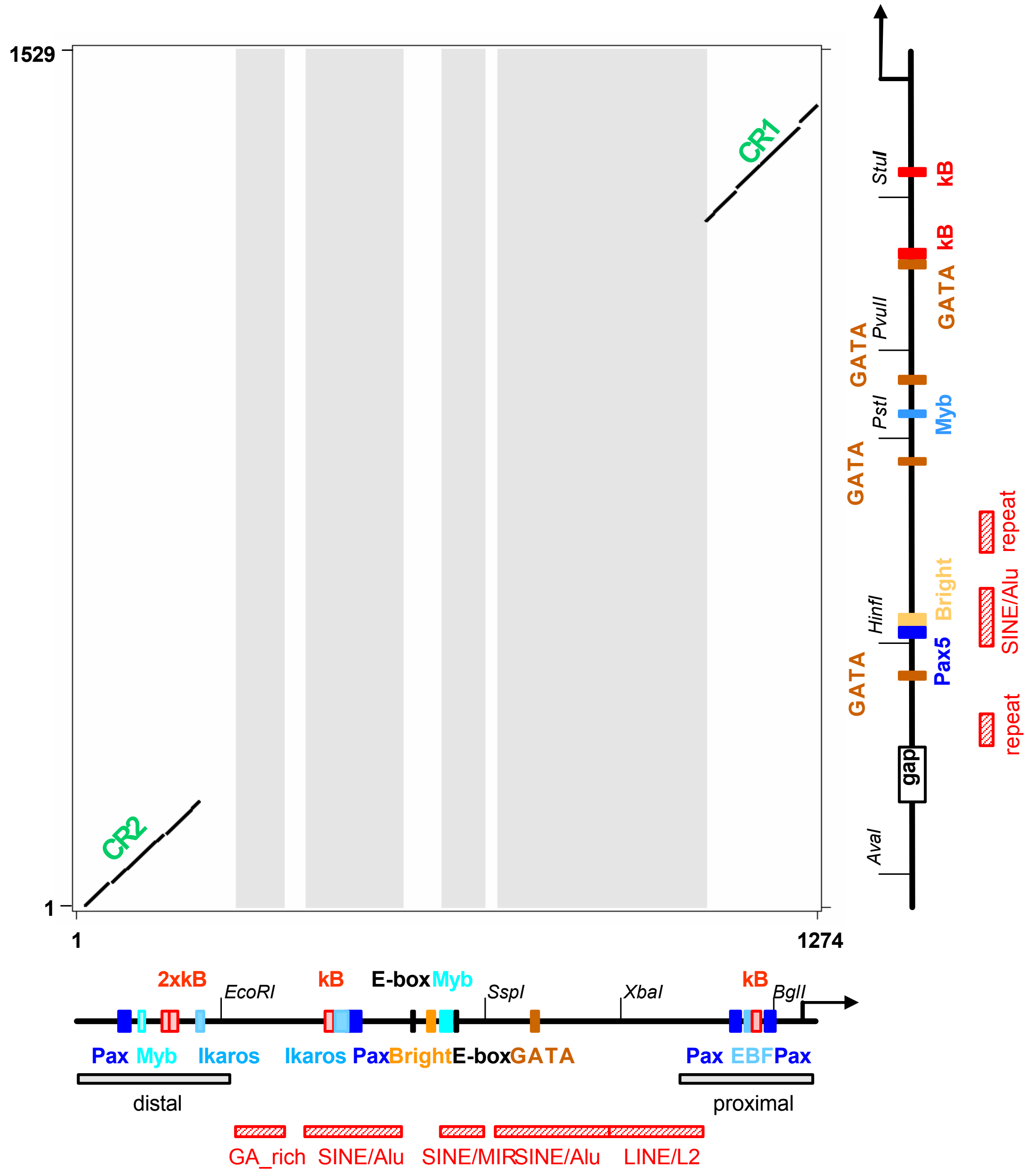
Schwartz, S., Zhang, Z., Frazer, K.A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R., and Miller, W. (2000). PipMaker: A Web Server for Aligning Two Genomic DNA Sequences. Genome Res. 10, 577-586.

U119 Experimental Oncology

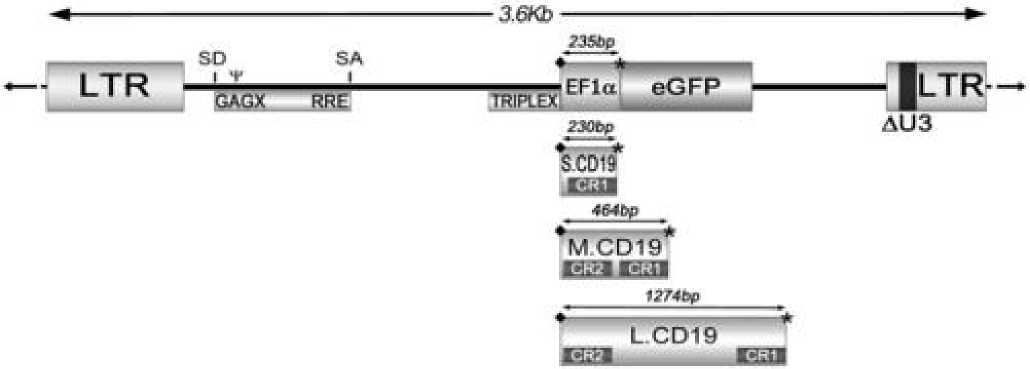
Selection of CD19 B-cell specific regulatory sequence and design of CD19-GFP lentiviral vector

Moreau, T., F. Bardin, J. Imbert, C. Chabannon, and C. Tonnelle. 2004. Restriction of transgene expression to the B-lymphoid progeny of human lentivirally transduced CD34+ cells. Mol.Ther,. 2004, in press.

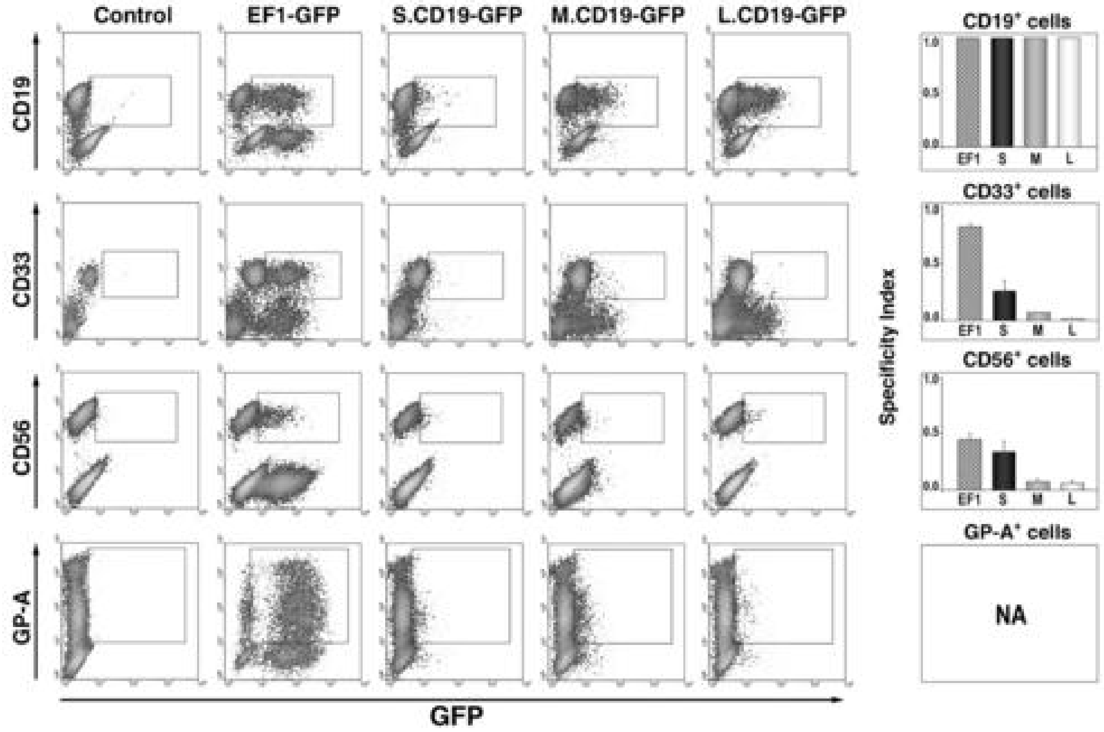
Computational analysis of homologies between Human and Mouse CD19 gene 5’ regions



A. Recombinant CD19-GFP lentiviral vectors



B. GFP expression in the progeny of transduced CD34$^+$ progenitor cells differentiated in vitro



From Moreau et al., Mol.Ther, 2004.

INSERM UMR599 Cancer Institute of Marseille

Director : Françoise Birg

Transcription Team

Michèle Algarté
Régis Costello
Bénédicte Delaval
Brigitte Kahn-Perlès
Patrick Lécine
Carol Lipcey
Pascal Rameil
Jung-Hua Yeh
Jean Imbert

Immunology Team

Chantal Cerdan
Jacques Nunès
Daniel Olive

Gene & Cellular Therapy Team

Thomas Moreau
Christian Chabannon
Cécile Tonnelle

Immunology Center of Marseille-Luminy

Salvatore Spicuglia
Sanjeev Kumar
Pierre Ferrier

INSERM U363, Paris

Fabrice Gouilleux

Curie Institute, Orsay

Jacques Ghysdael

ISREC, Lausanne

Philip Bücher
Markus Nabholz

ICRF, London

Carol Beadling
Doreen Cantrell

Genzentrum, Martinsried,

Patrick Baeuerle