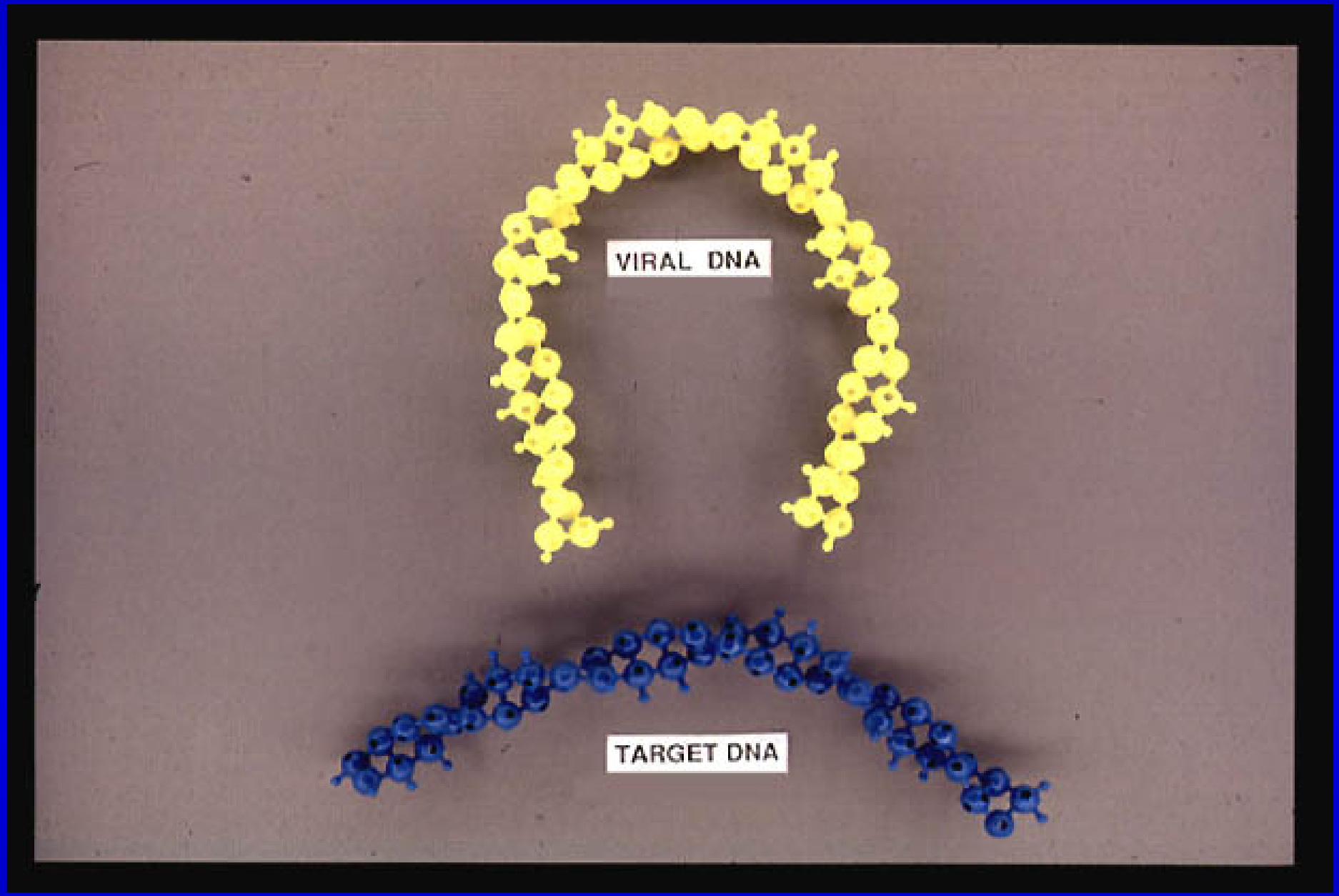
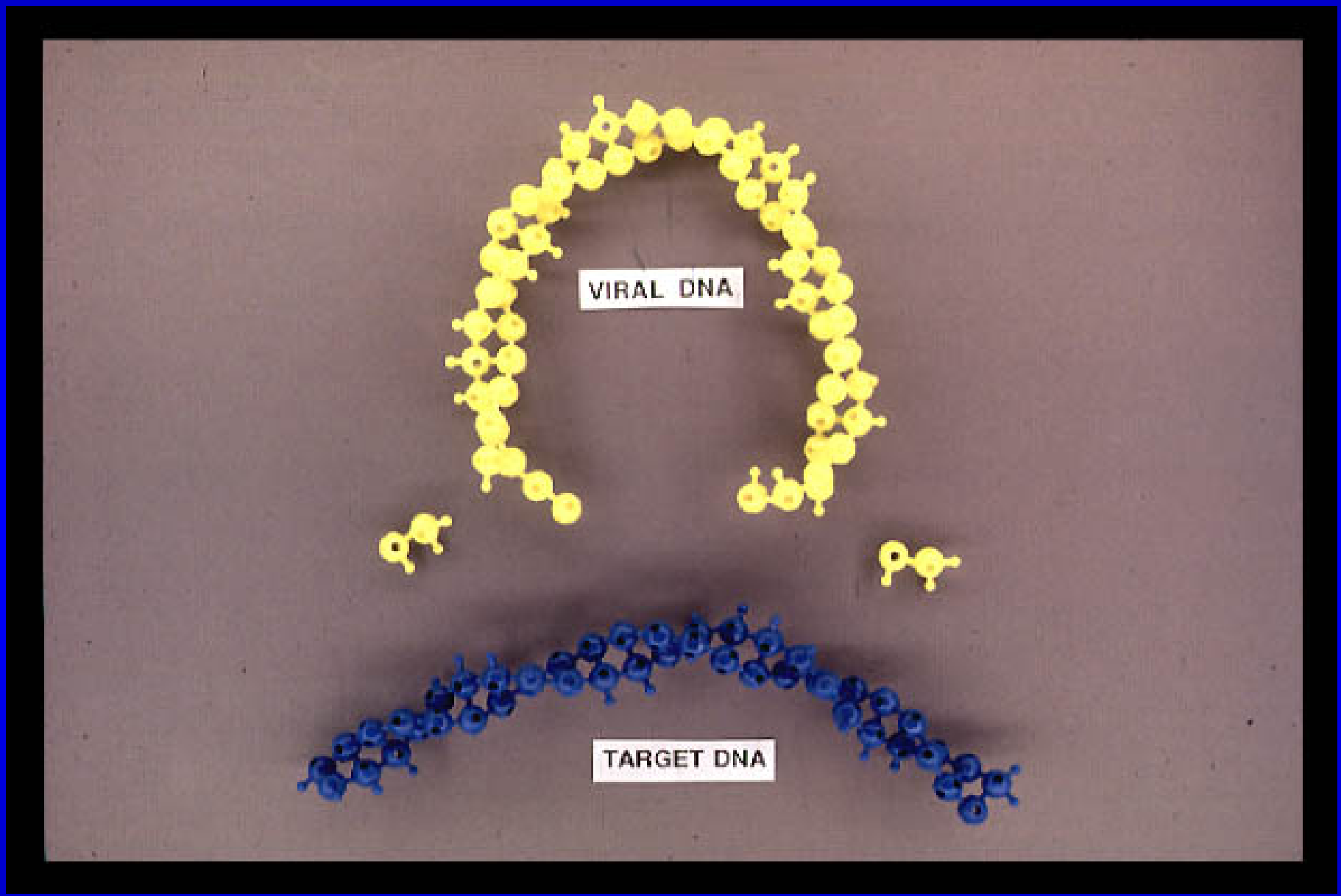
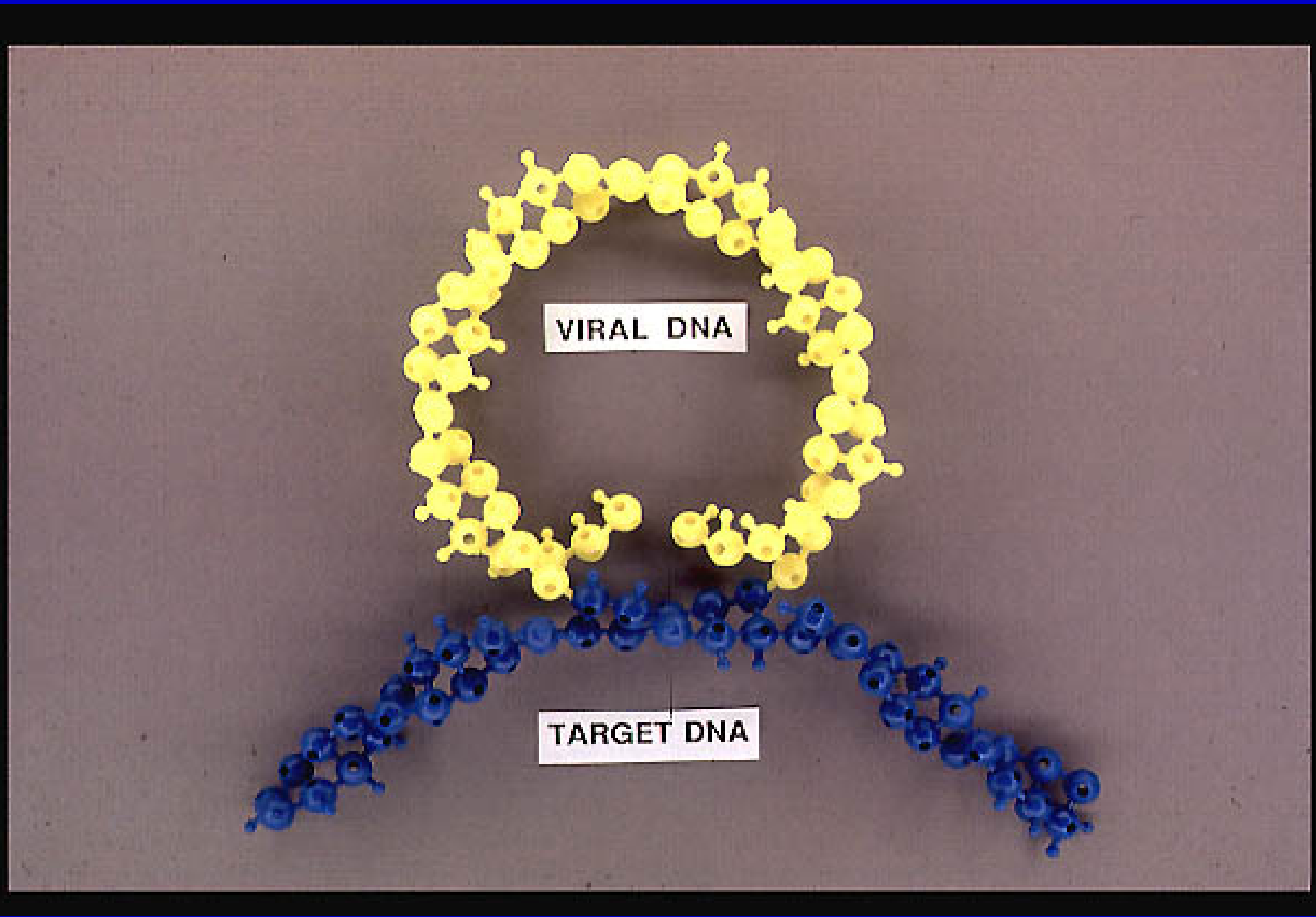
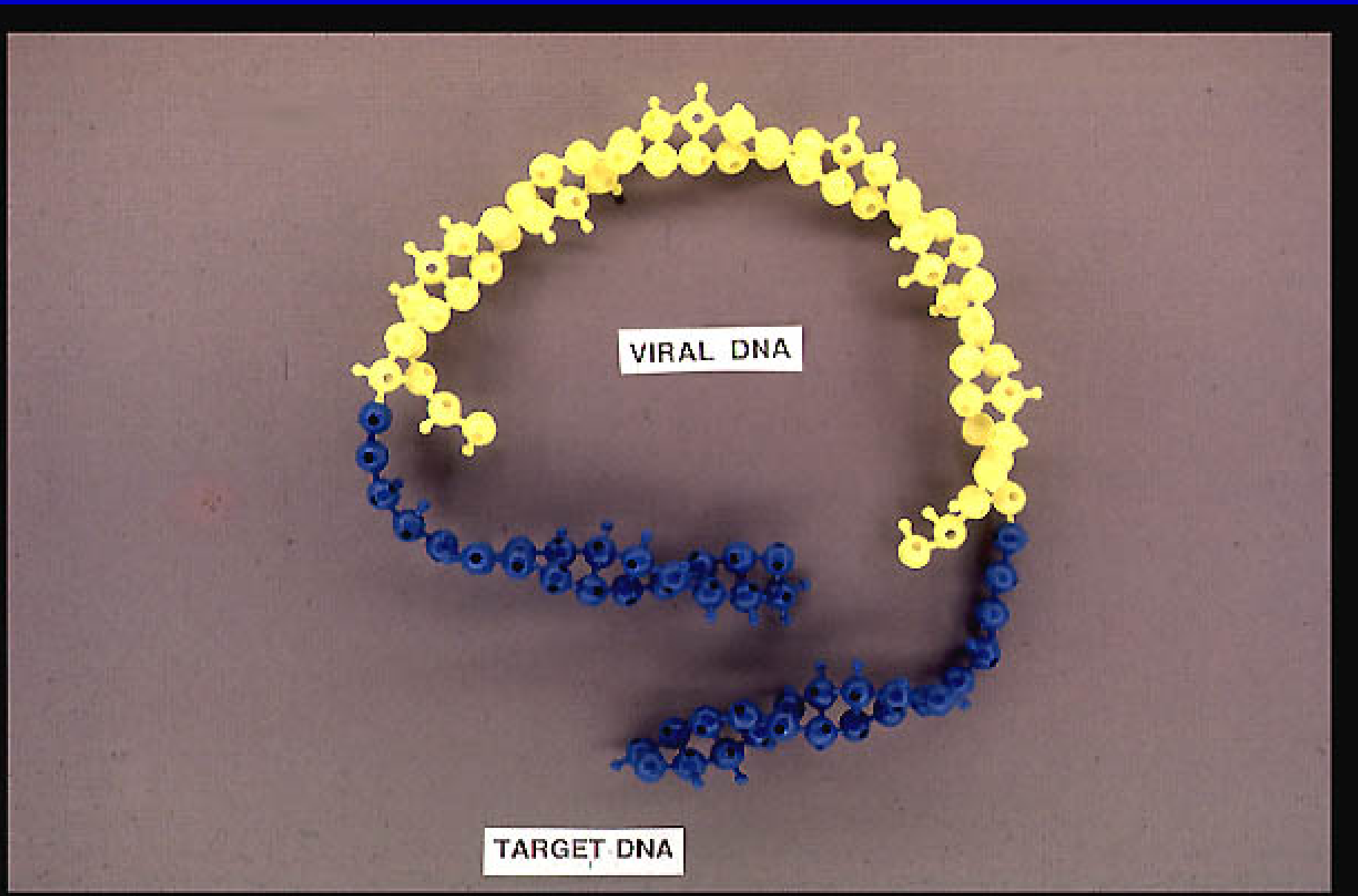
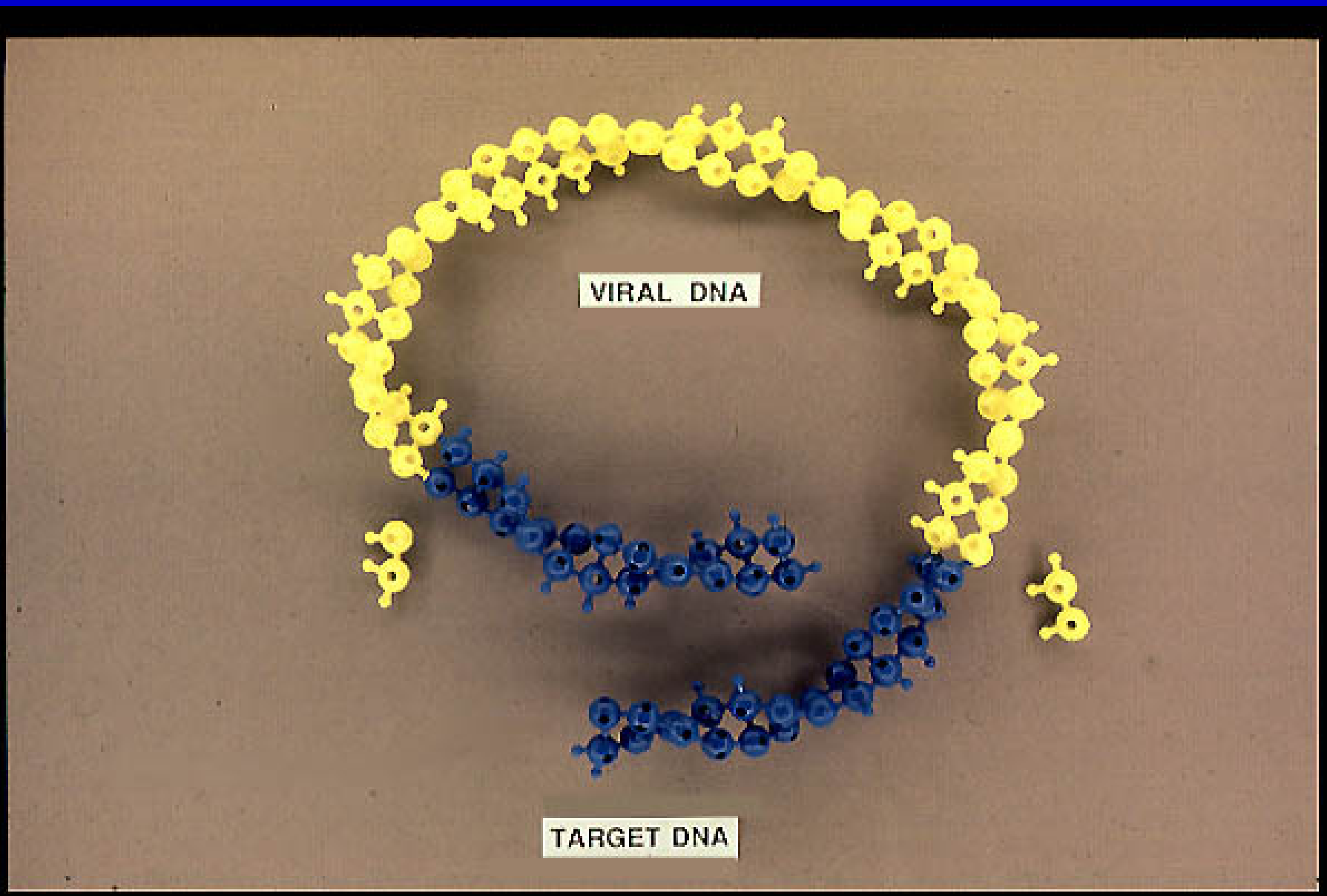
- HIV integration targeting
- Retroviral integration targeting
- Determinants of integration targeting


VIRAL DNA
TARGET DNA

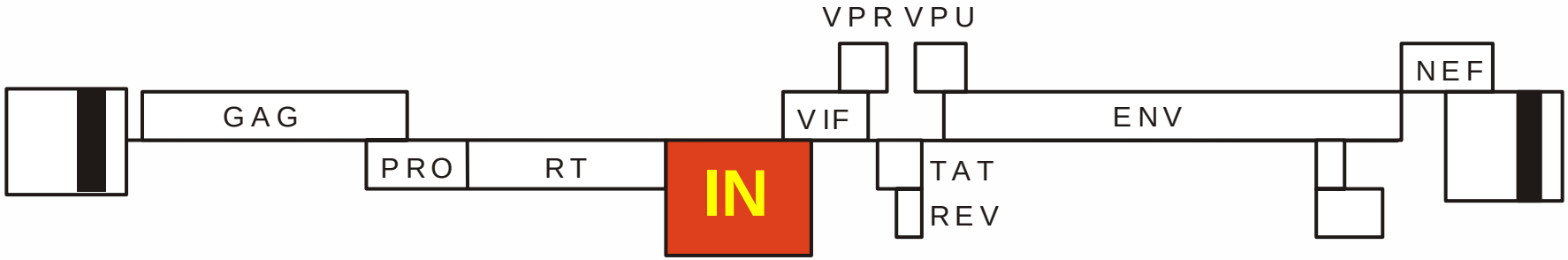

VIRAL DNA
TARGET DNA


VIRAL DNA
TARGET DNA


VIRAL DNA
TARGET DNA


VIRAL DNA
TARGET DNA

The integrase coding region



Genome-wide survey of integration targeting

Integration in vivo

- Infect cells with HIV or vector based on HIV, ASLV or MLV
- After 48 hours, isolate DNA
- Cleave DNA with restriction enzyme
- Ligate DNA linkers onto cleaved DNA ends
- Amplify by PCR with one primer in HIV and another that binds the linker
- Clone PCR fragments
- Automated colony preparation and sequencing (over 5000 sites)
- Map flanking human DNA segments on the human genome sequence
- Bioinformatic and statistical analysis (compare to random sites, etc)

Schroder, Shinn, Chen, Berry, Ecker and Bushman,2002 Cell 110, p.521-529
Wu, Crise and Burgess, 2003 Science 300, p. 1749-1751
Mitchell, Beitzel, Schroder, Shinn, Chen, Berry, Ecker and Bushman, PLoS Biology, 2004

Sequences of over 3000 integration sites

Schroder, Shinn, Chen, Berry, Ecker and Bushman,2002 Cell 110, p.521-529
Wu, Crise and Burgess, 2003 Science 300, p. 1749-1751
Mitchell, Beitzel, Schroder, Shinn, Chen, Berry, Ecker and Bushman, PLoS Biology, 2004

HIV integration frequency and genes

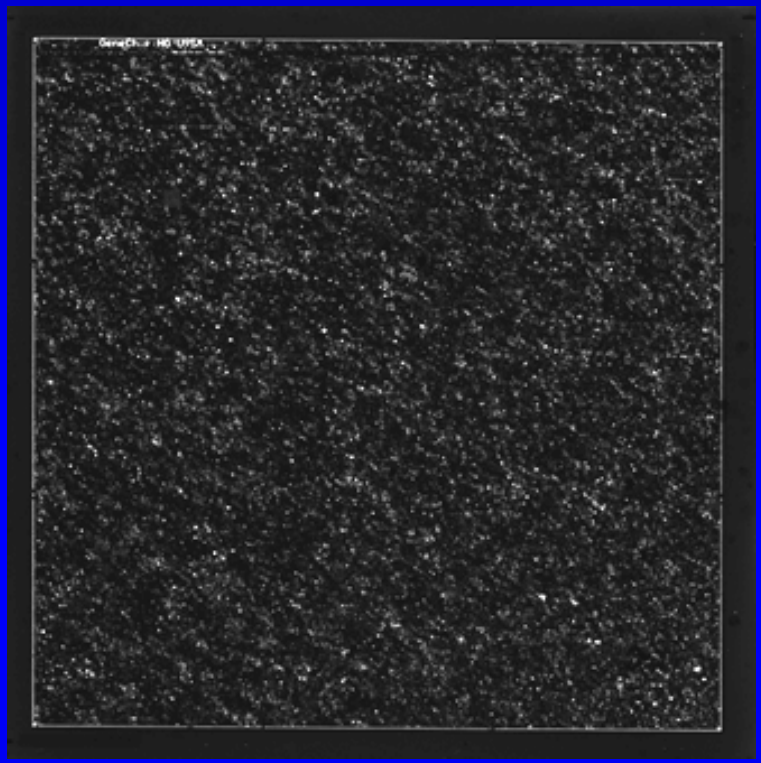
Frequency of integration in genes

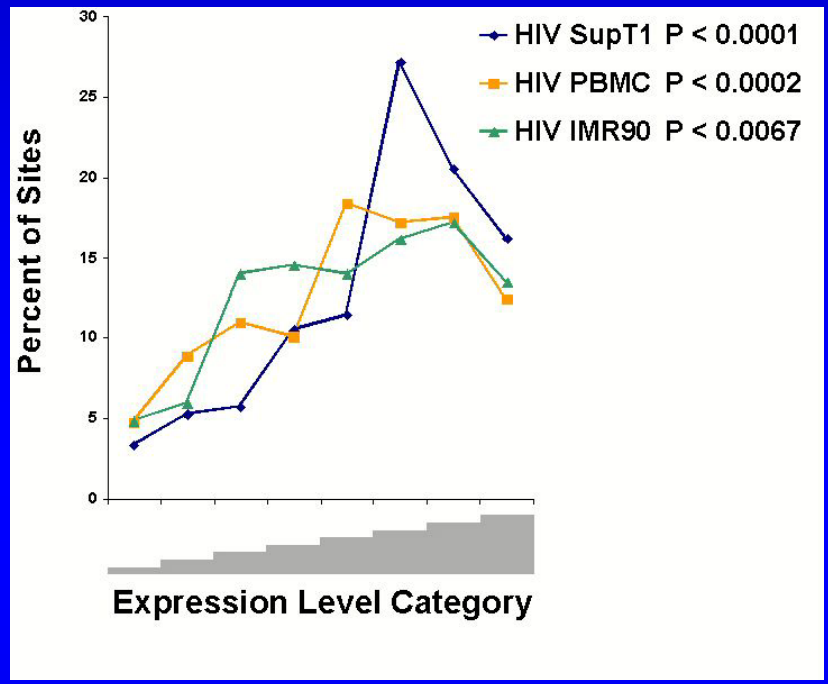
HIV	Sup T1	69%
HIV	PBMC	73%
HIV	IMR90	69%
Total genome		~35%

HIV integration strongly favored in genes ($p<0.0001$)

Conclusions hold after correction for restriction site placement in the human genome

HIV integration frequency and gene activity





Conclusions

•Median average difference around 2-fold higher for genes that hosted integration events than for whole chip--thus gene activity favors integration.

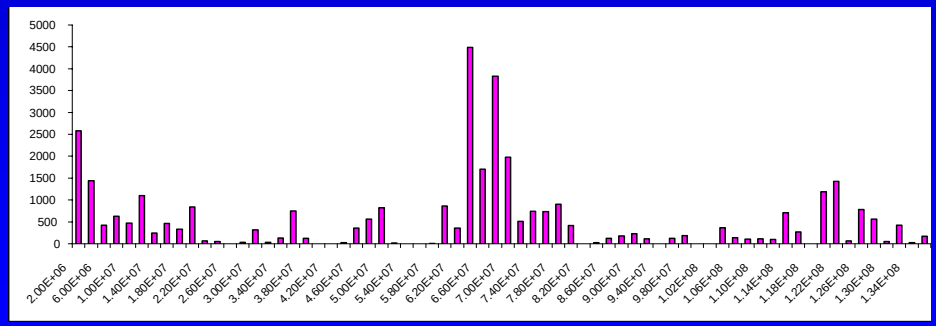
•Active genes strongly favored for HIV integration ($p<0.0001$), though highest expression category more weakly expressed.

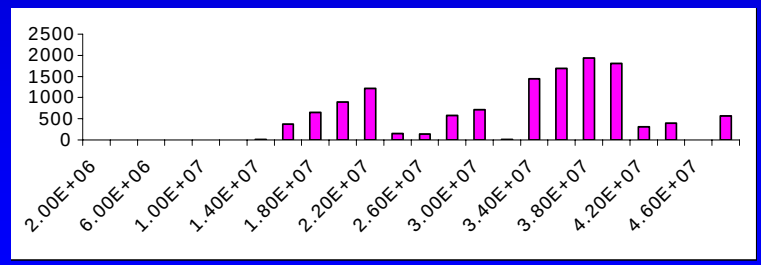
•Modest though significant effect of tissue-specific transcription

Transcriptional Intensity Versus Integration Intensity

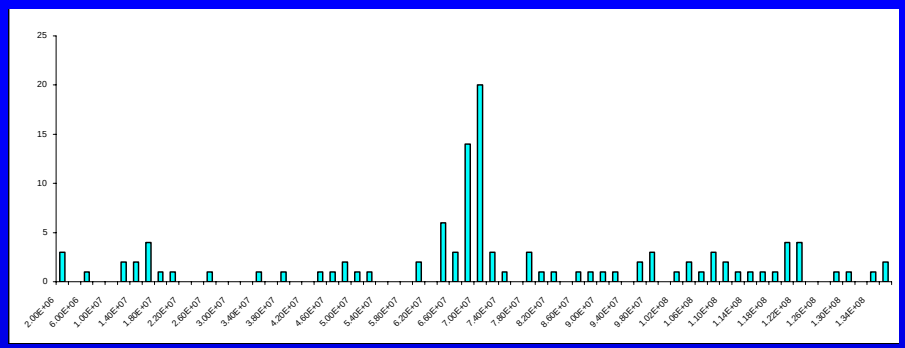
Chromosome 11 Chromosome 22

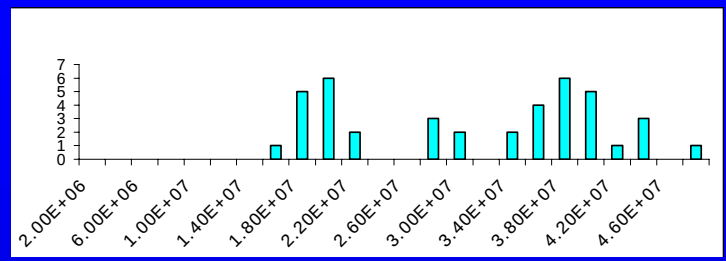
Transcriptional Intensity (EST data from Sanger Center)





Integration Intensity





Suggests gene-rich chromosomal domains are favored for integration.
Favorable and unfavorable regions interleaved in gene-rich regions.

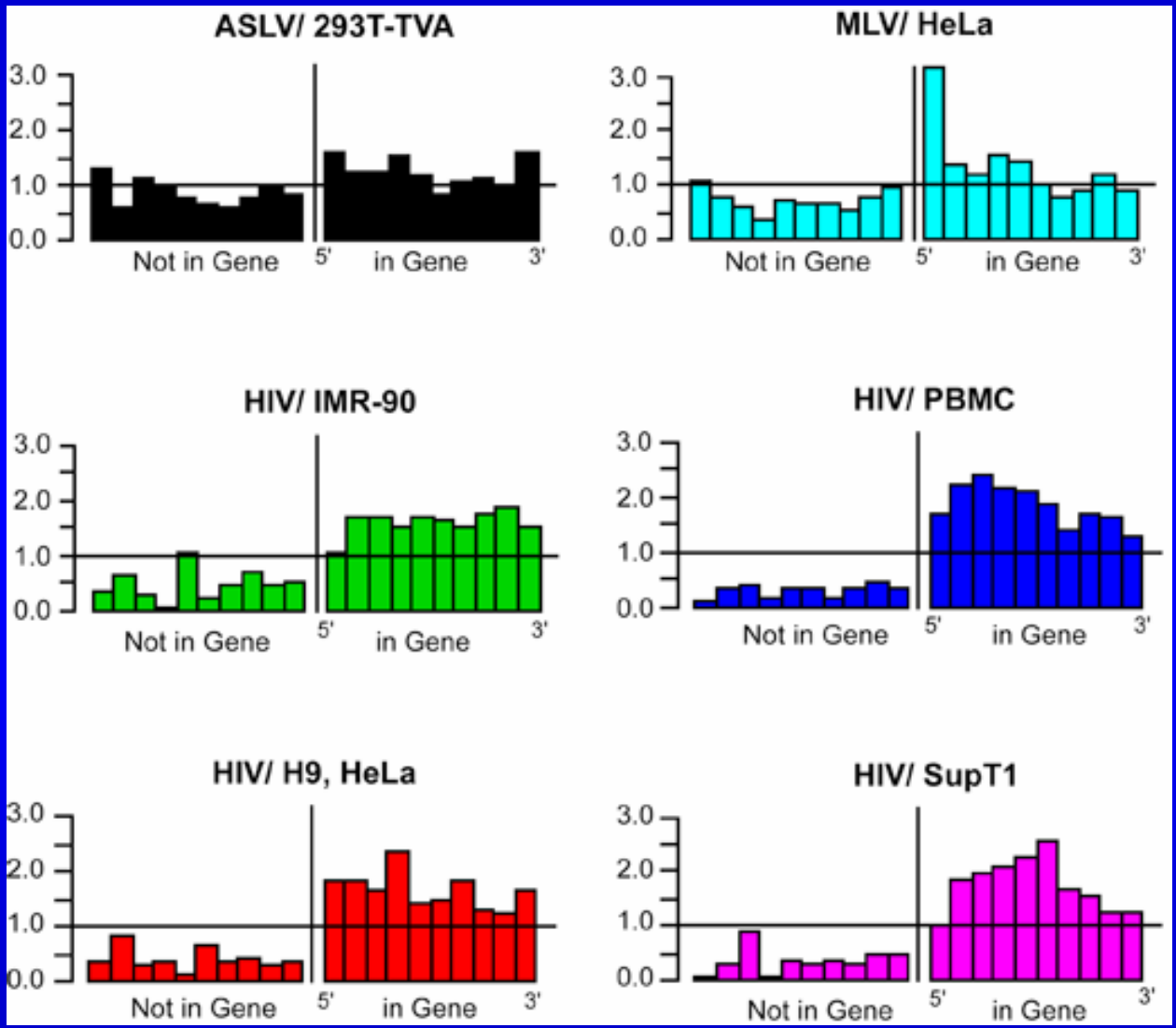
Mitchell et al., PLoS Biology 2004

Conclusions: Target site selection by HIV integration complexes

- HIV favors integration in active genes
- Seen in primary cells and cell lines
- Clear though modest influence of tissue specific transcription
- Integration sites cluster
- Favored regions small in size (100-250 kb)--length of one or a few genes. For HIV, interspersed with unfavorable CpG islands.
- Role in HIV life cycle?

- HIV integration targeting
- Retroviral integration targeting
- Determinants of integration targeting

Integration site selection by HIV, MLV, and ASLV



ASLV, MLV, and HIV all significantly different from one another

MLV favors integration near transcription start sites (Burgess and coworkers)

First two adverse events: MLV-based vector integrated in this unfavorable window

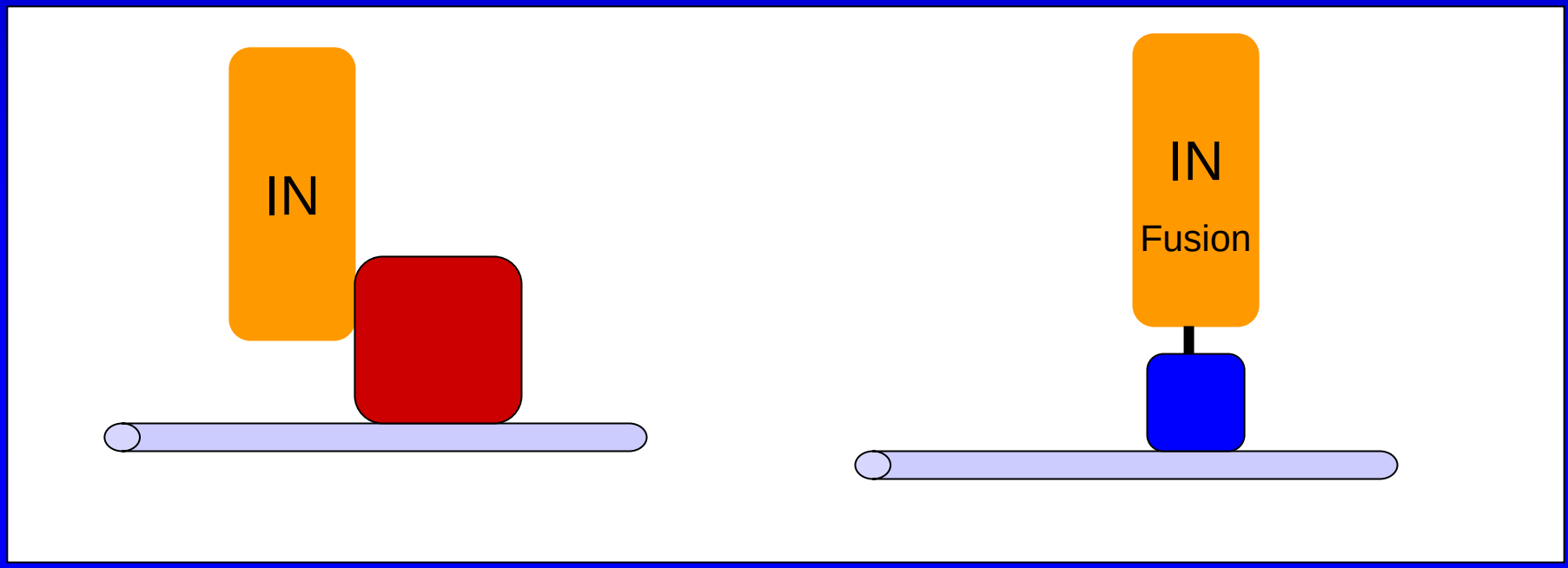
MLV and HIV/H9,Hela data from Burgess and coworkers (Science 2003)

•Target site selection by ASLV and MLV integration complexes

•MLV favors integration in gene transcription start regions and CpG islands (Burgess)
•MLV more effective at gene trapping than HIV (Naldini and coworkers, 2005)
•ASLV shows weaker preference for active genes
•Targeting by ASLV potentially favorable for human gene therapy
•All retroviruses so far studied different from each other. Mechanisms?

- HIV integration targeting
- Retroviral integration targeting
- Determinants of integration targeting

Controlling integration targeting?



Natural targeting by IN-binding to cellular DNA binding proteins

Mechanism used by yeast Ty elements
Mechanism for retroviruses as well?

Yeast work: studies from Voytas, Sandemeyer, Boeke, Levin and coworkers.

Artificial tethering by retroviral integrase enzymes

Works great in vitro
So far not effective in vivo
Interference by natural tethering mechanism?

Bushman (1994) PNAS 91 9233-9237
Bushman (1995) Science 267, 1443-1444.
Related studies from S. Chow; Katz and Skalka

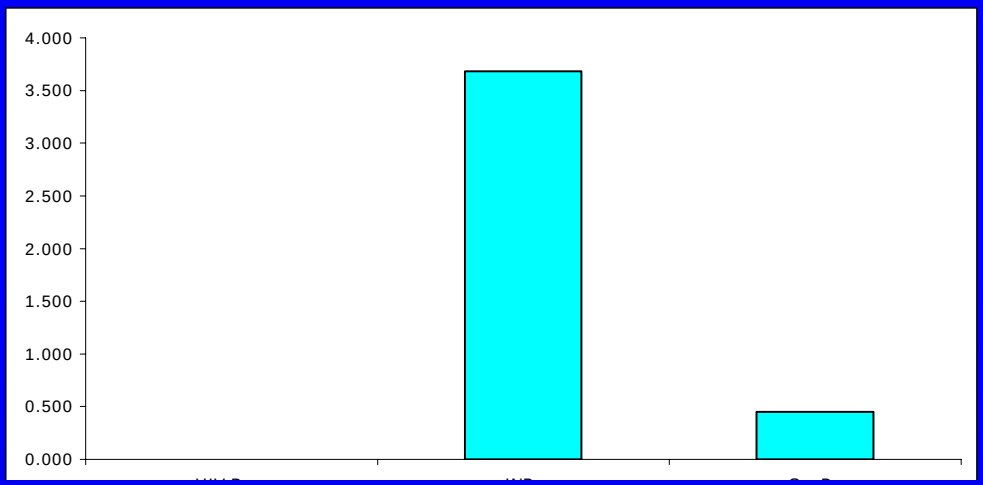
Ongoing Studies

Viral determinants of integration targeting
Compare integration targeting in chimeras of HIV and MLV (from Emerman and colleagues)

Cellular determinants of integration targeting
Study integration targeting in cells knocked down or deleted for candidate targeting factors.

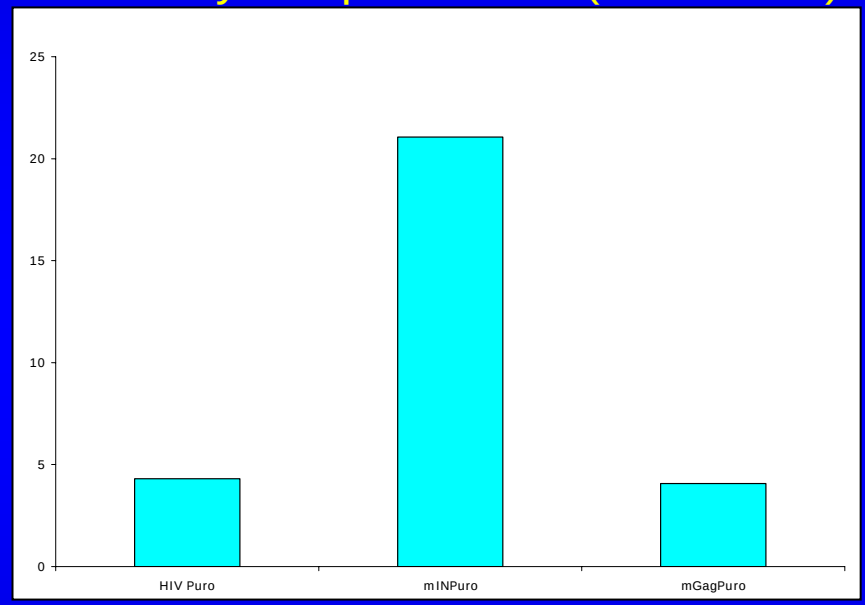
mINPuro Resembles MLV in favoring integration near transcription start sites and CpG islands

Proximity to RefSeq Start sites (+ or - 2.5 Kb)



HIVPuro mINPuro mGagPuro

Proximity to CpG Islands (+ or - 5 Kb)



HIVPuro mINPuro mGagPuro

mINPuro versus HIVPuro achieves P=0.005

Comparison of HIVPuro to mINPuro achieves P=0.0004

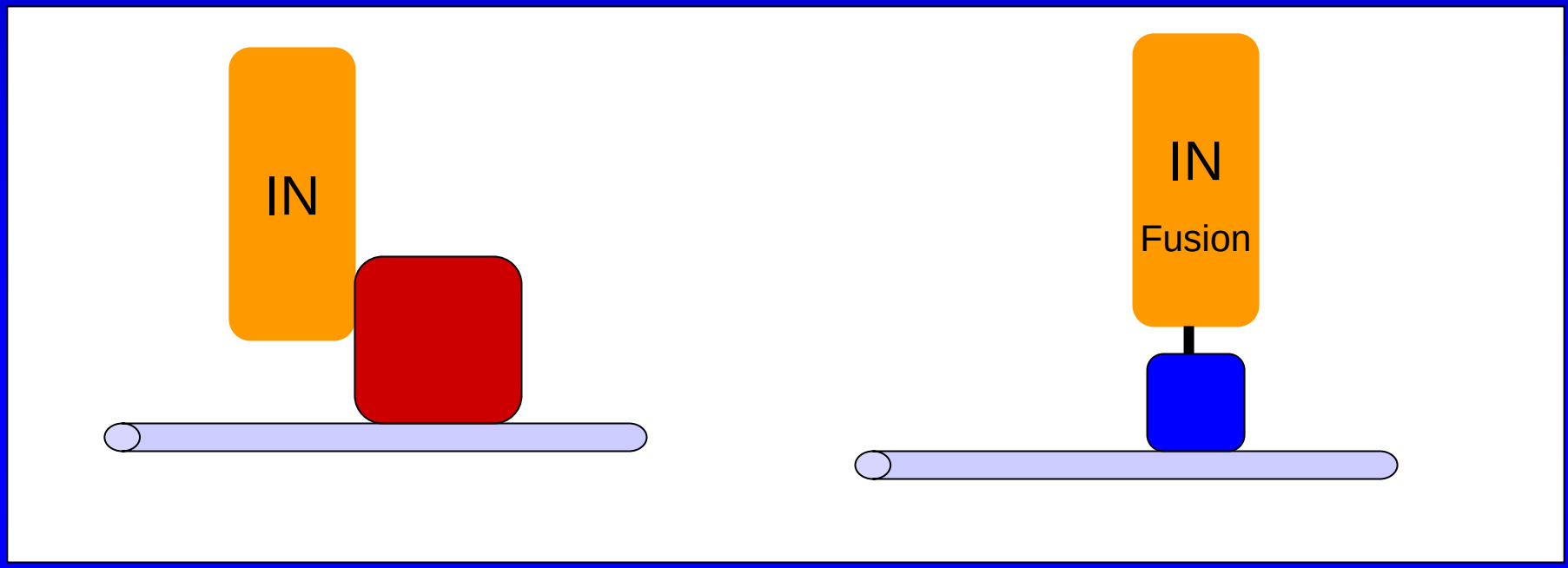
Conclusions

- mINHIV favors integration near transcription start sites and CpG islands (as seen in studies of wt MLV). Therefore MLV IN confers these targeting properties.
- Exact opposite of what you would like for gene therapy (but of course, the represents an early step in modulating specificity).
- Studies of mutant cells lines--specify tethering factors for HIV

Targeting in other integrating systems

- SIV. Studied in primate models; resembles HIV.
- AAV. Strongly favors integration at transcription start sites and CpG islands. Much expression from unintegrated DNA. Often rearranges DNA at target sites.
- LINE elements. Conflicting data on favored integration in genes. Often rearranges DNA at target sites.
- DNA transposons (sleeping beauty, mariner, etc.). Weak bias in favor of integration in genes. Involves enzyme that introduces DNA ds breaks.
- Homologous recombination. Allows integration at predetermined sites. Efficiency major obstacle.

Strategies for controlling integration targeting



- Switch integration systems (ASLV, others?). Note: random integration should be 33% in transcription units
- Genetically mutate integrases to remove undesireable tethering interactions.
- Engineer in new tethering interactions. Success in yeast models (Voytas and coworkers). However, concentration of nonspecific competitor DNA in nucleus big challenge.

- HIV integration targeting
- Retroviral integration targeting
- Determinants of integration targeting

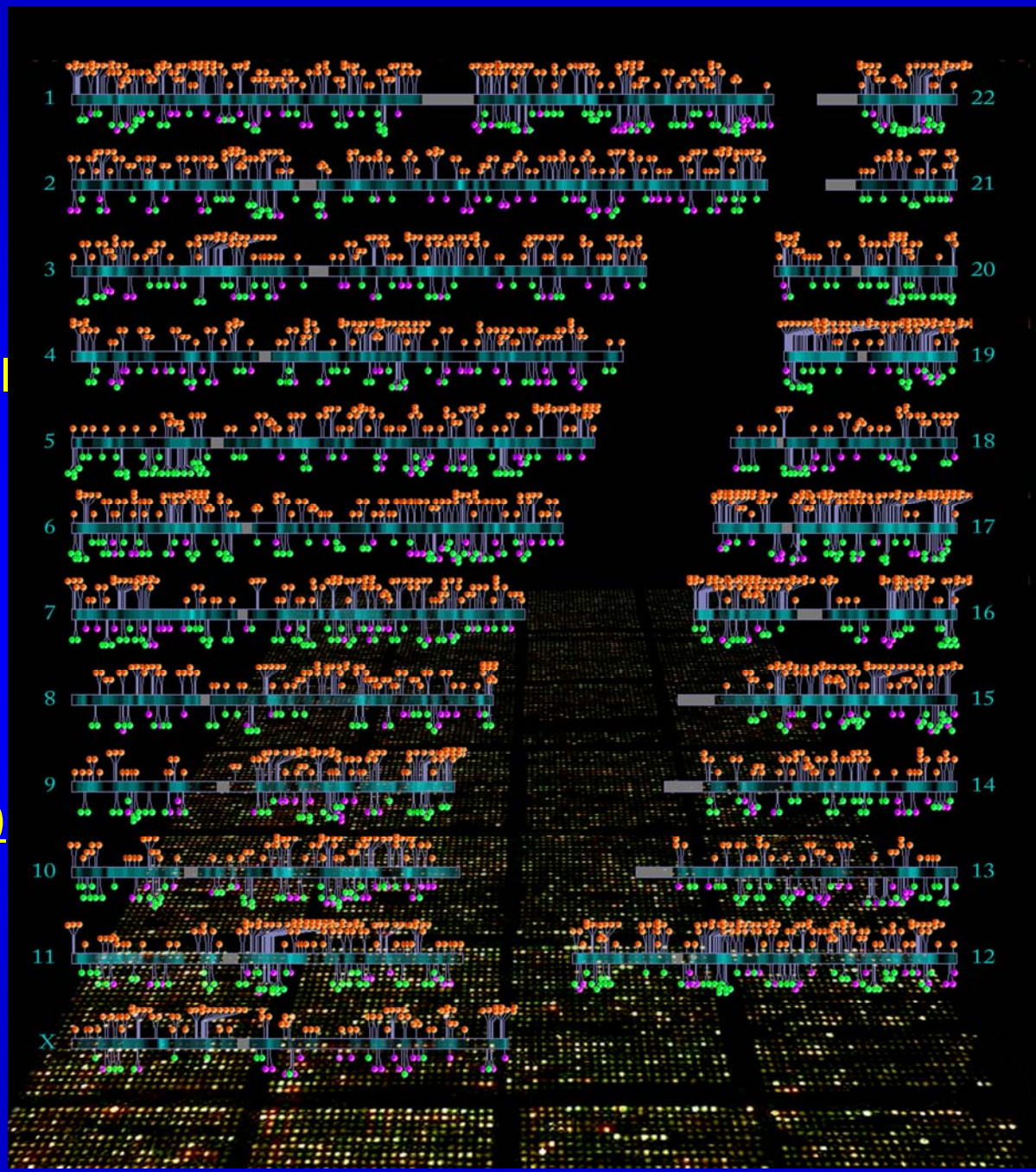
CREDITS

Bushman Laboratory
Astrid Schroder, Rick Mitchell
Brett Beitzel, Mary Lewinski,
Steve Barr, Jeremy Leipzig

Ecker Laboratory (Salk)
Paul Shinn, Huaming Chen,
Joe Ecker

Emerman Laboratory (Hutch)
Mike Emerman, Masahiro
Yamashita

Chuck Berry (UCSD)



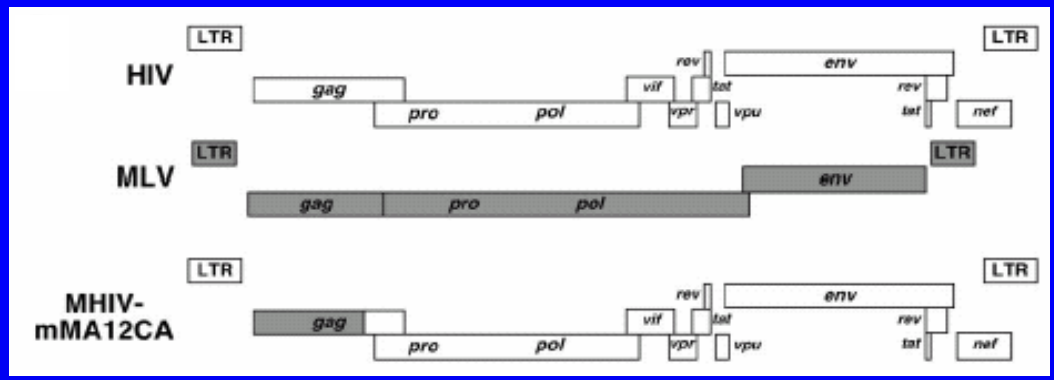
Integration targeting by HIV derivatives containing portions of MLV

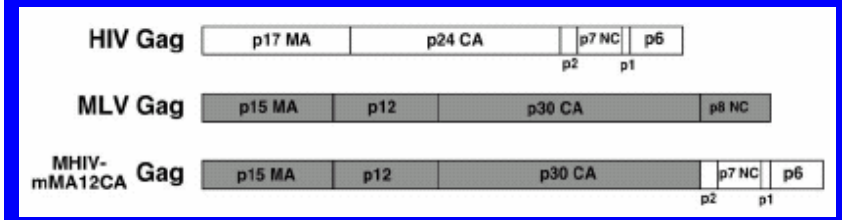
Capsid Is a Dominant Determinant of Retrovirus Infectivity in Nondividing Cells

Masahiro Yamashita and Michael Emerman*

Divisions of Human Biology and Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109-1024

JOURNAL OF VIROLOGY, June 2004, p. 5670–5678

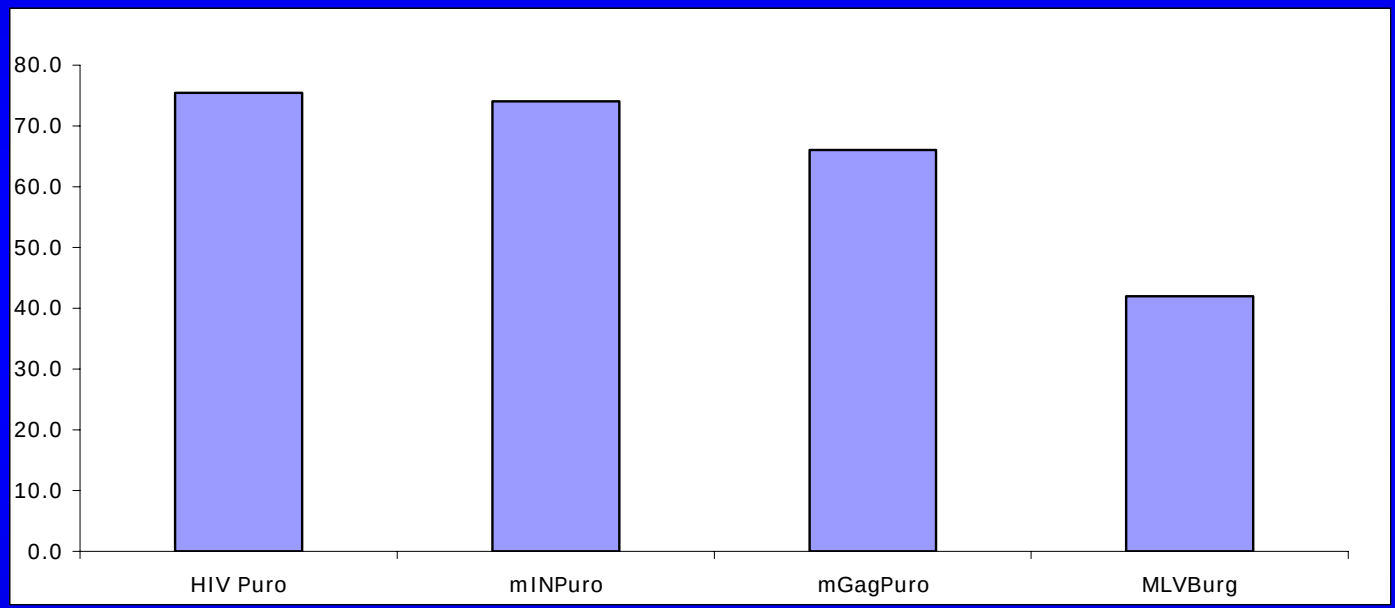




- Chimeric HIV with MLV MA, p12 and CA replacing most of HIV gag: cannot efficiently infect nondividing cells (Yamashita and Emerman, JV 2004)
- Chimeric HIV with MLV IN. Shows correct MLV 4 bp target sequencing duplication (Yamashita and Emerman, unpublished)

Integration Targeting of Transcription Units

Data so far: HIVPuro: 227 sites, mINPuro: 193 sites, mGAGPuro: 436 sites

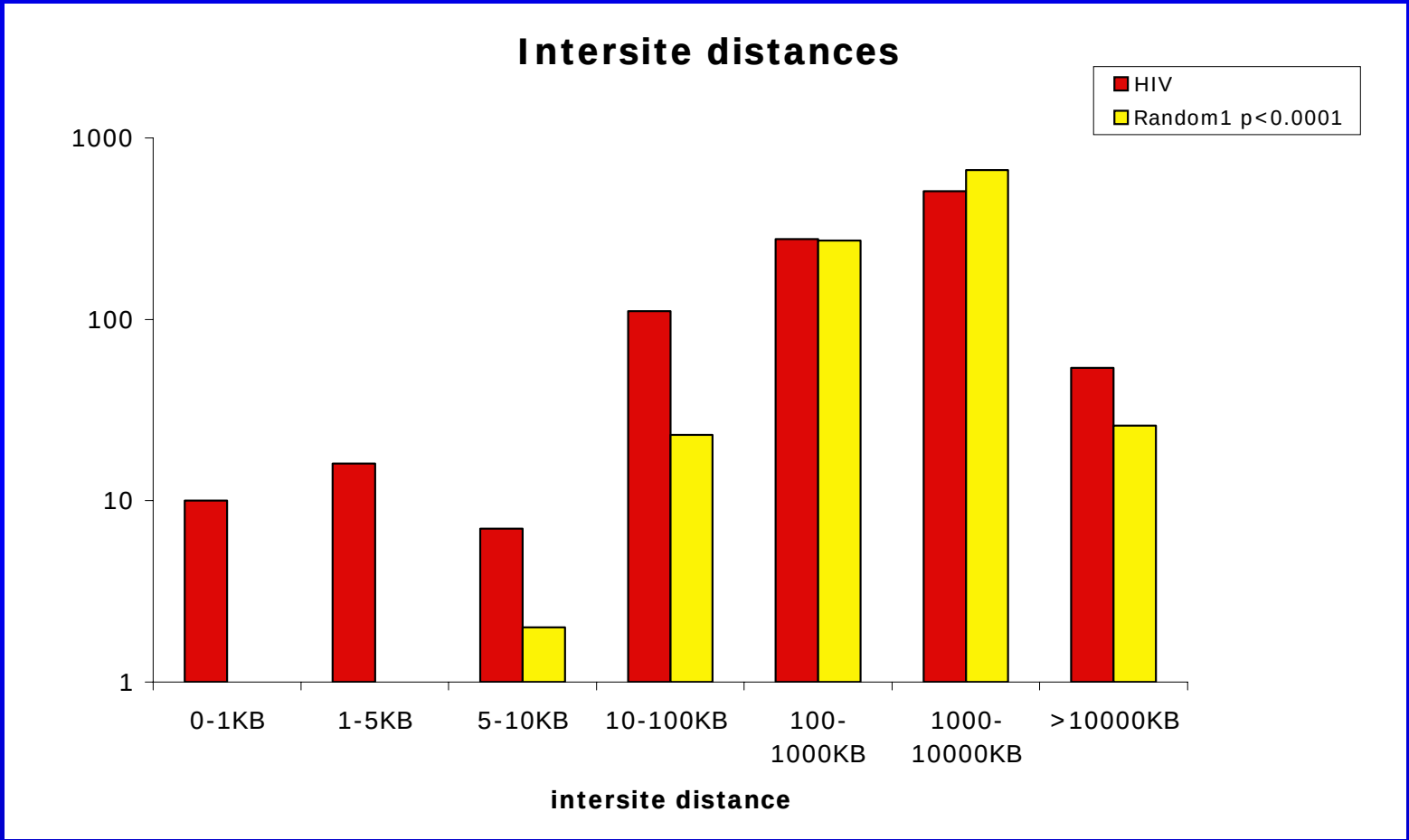


Statistics so far:
mINPuro versus HIVPuro: P=0.738
mGagPuro versus HIVPuro: P=0.0078

Lewinski, Leipzig, Shinn, Yamashita, Emerman, Ecker, and Bushman, unpublished

Regional hotspots for integration defined by multiple nearby hits

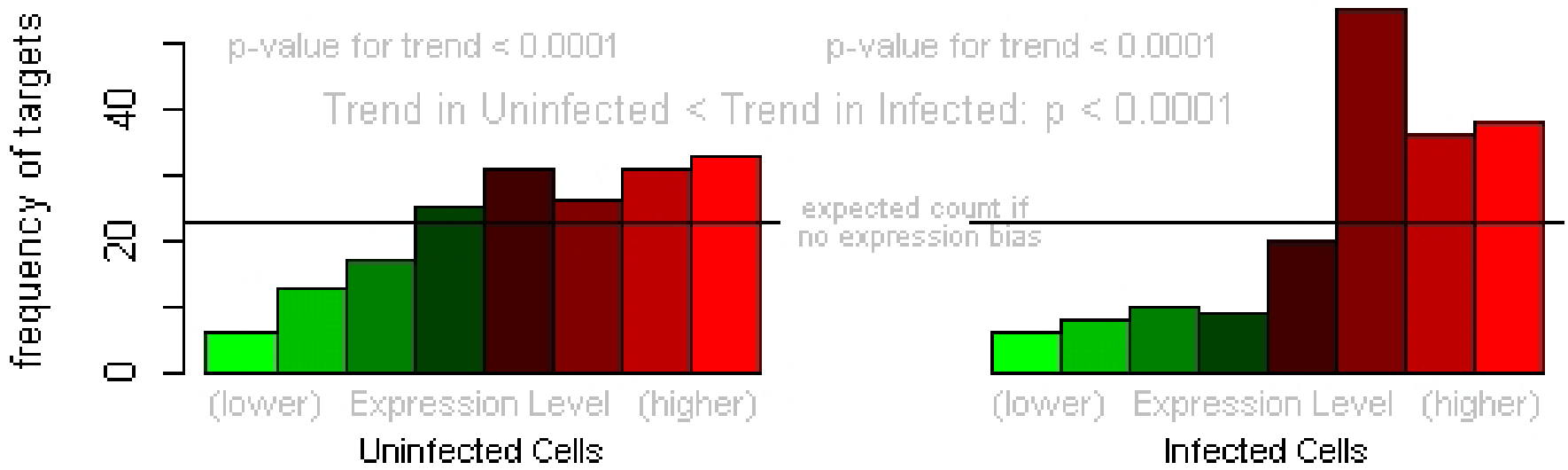
Analysis of intersite distances for 1012 integration events in leukocytes
Compare to the same number of calculated random sites distributed around the genome



Clustered integration sites define regions that are special in some sense
Cold regions are also seen

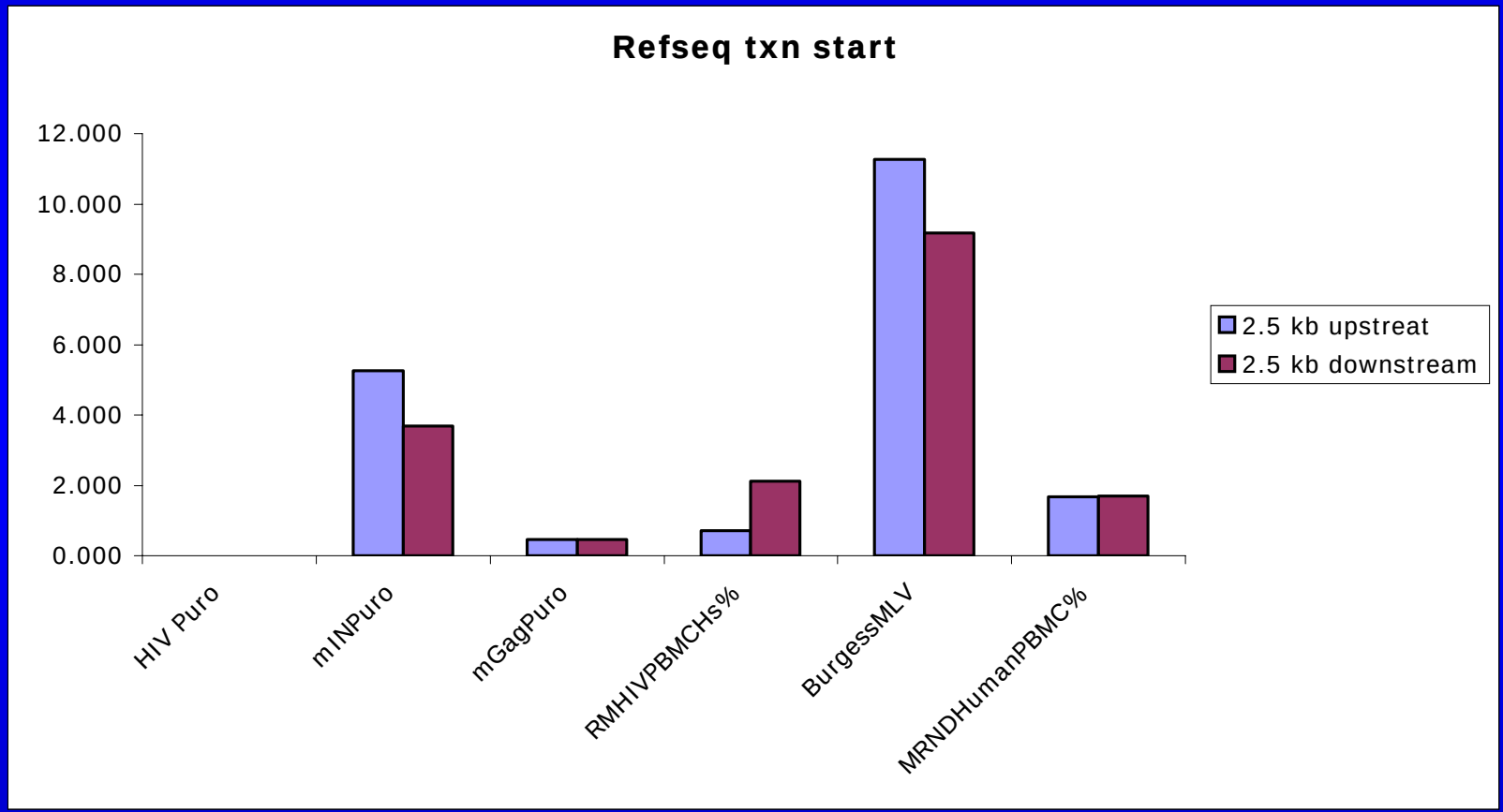
Correlation between gene expression and integration targeting

- A significant correlation is seen between targeting and transcription in cells before infection.
- Integration is particularly favored in genes active after infection.



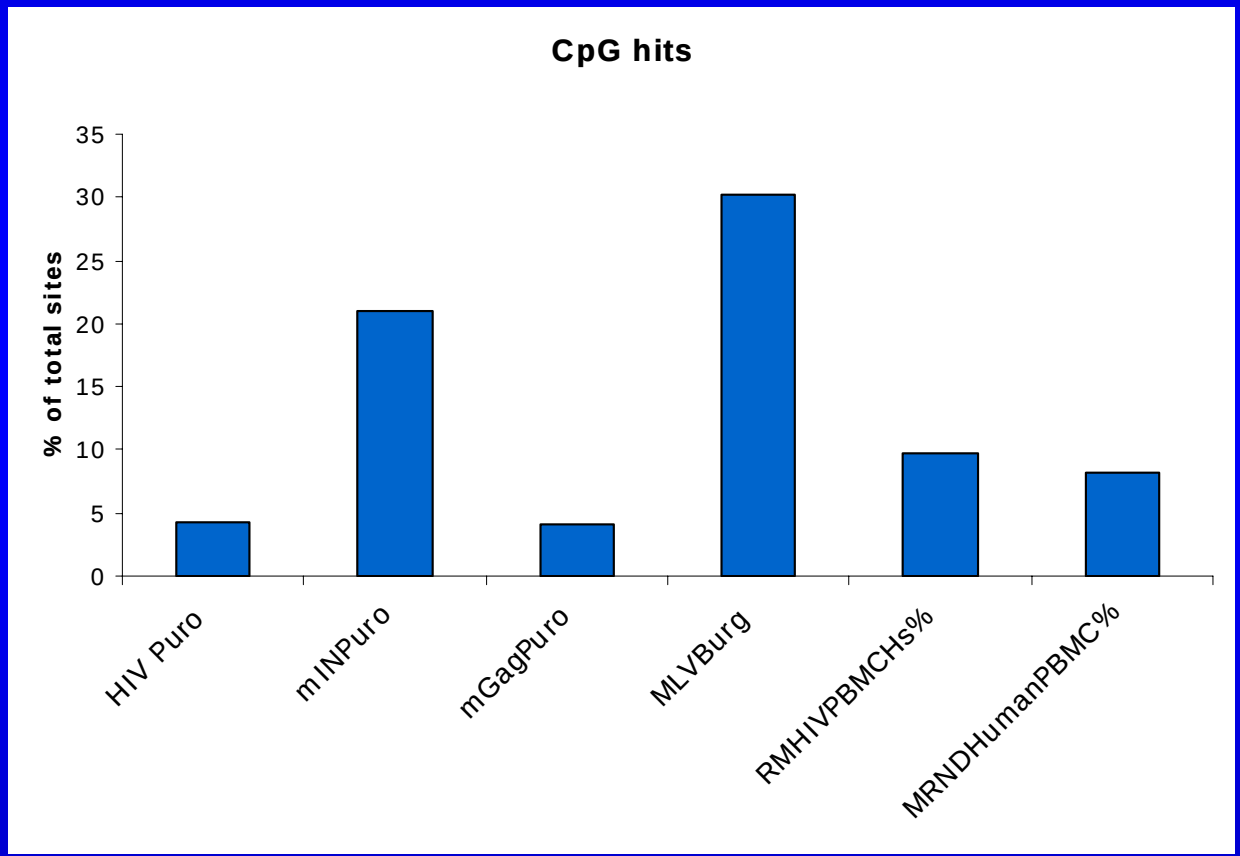
Quantitative analysis by Charles Berry

Integration near transcription start sites



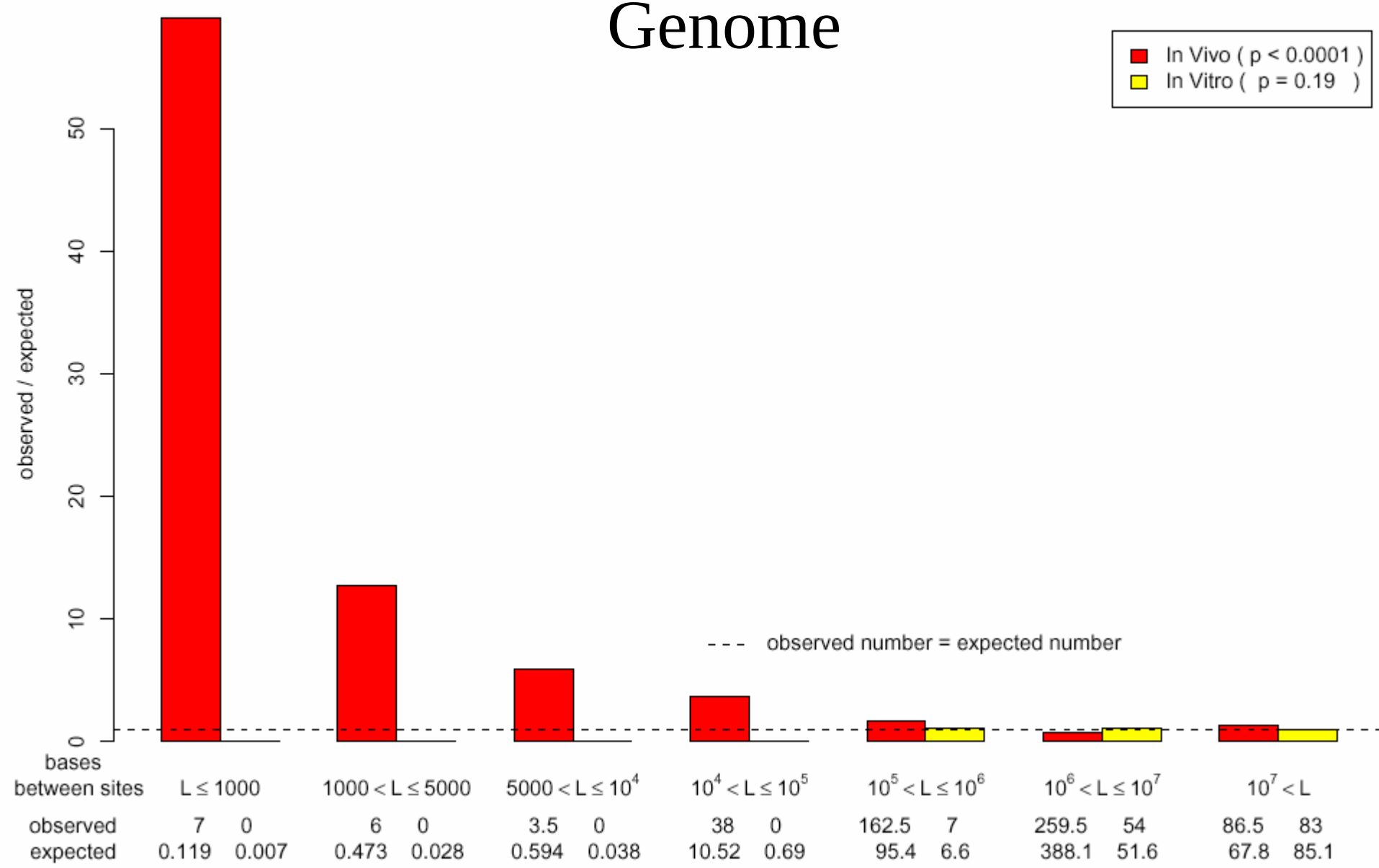
mINPuro versus HIVPuro achieves P=0.005

Integration Near CpG Islands



Comparison of HIVPuro to mINPuro achieves P=0.0004

Integration Site Spacing on the Human Genome



Chromosomal Features Associated with HIV-1 Integration Sites

Chromosomal feature	% in human genome	% at in vivo integration sites	% at in vitro integration sites
Transcription units	~33%[#]	69% (p<<0.0001)	35% (p=0.76)
SINES			
Alu	10.6%	15.9% (p=0.001)	13.2% (p=0.55)
MIR	2.2%	0.7% (p=0.03)	0.8% (p=1.00)
DNA elements	2.8%	2.2% (p=0.46)	0.8% (p=0.52)
LTR elements (HERV)	8.3%	3.7% (p=0.0002)	6.6% (p=0.24)
LINE	20%	17.0% (p=0.10)	16.5% (p=1.0)
Satellite			
alpha Satellite	UN	0.4%	1.7%
beta Satellite	UN	0%	1.7%

The integration sites studied included those mapped to unique locations on the genome and those in identifiable repeats. p values are for comparison of each integration site population to the human genome.
# = estimated
UN=unknown

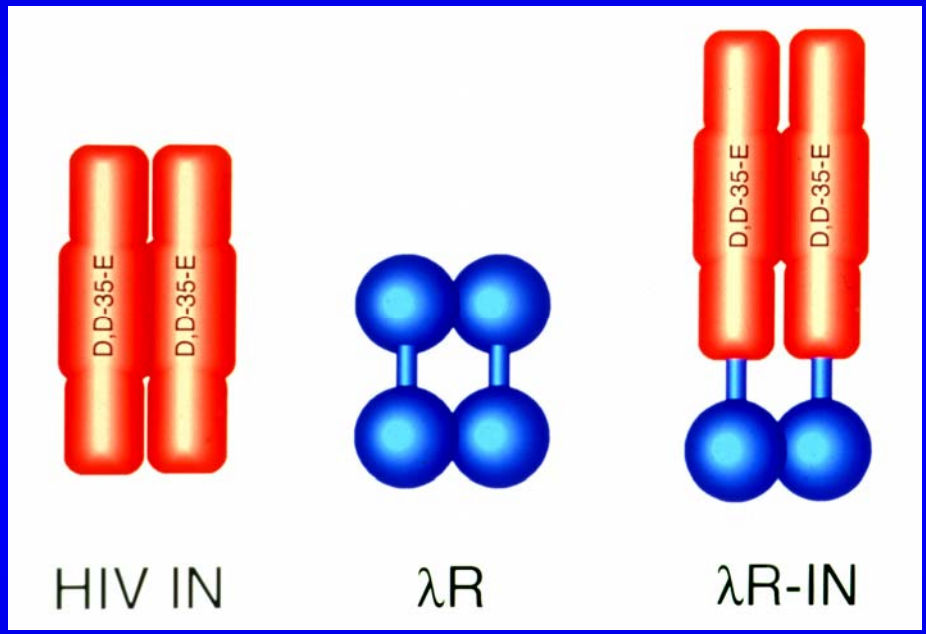
HERV repeats and HIV integration sites negatively correlated
Note that the in vitro control doesn't differ from genome as a whole: important validation of methods

Favored integration: removing something bad or adding something good?

Repression at disfavored sites?
P element analogy

Carteau et al., (1999) JV
Schroder et al. (2003) Cell
For review and additional references, see:
Bushman (2003) Cell 115, 2.

Tethering to favored sites?
Yeast retroelement paradigm



Bushman (1994) PNAS 91 9233-9237
Bushman (1995) Science 267, 1443-1444.
Bushman and Miller (1997) J. Virol. 71, 458-464.
Related studies from S. Chow; Katz and Skalka