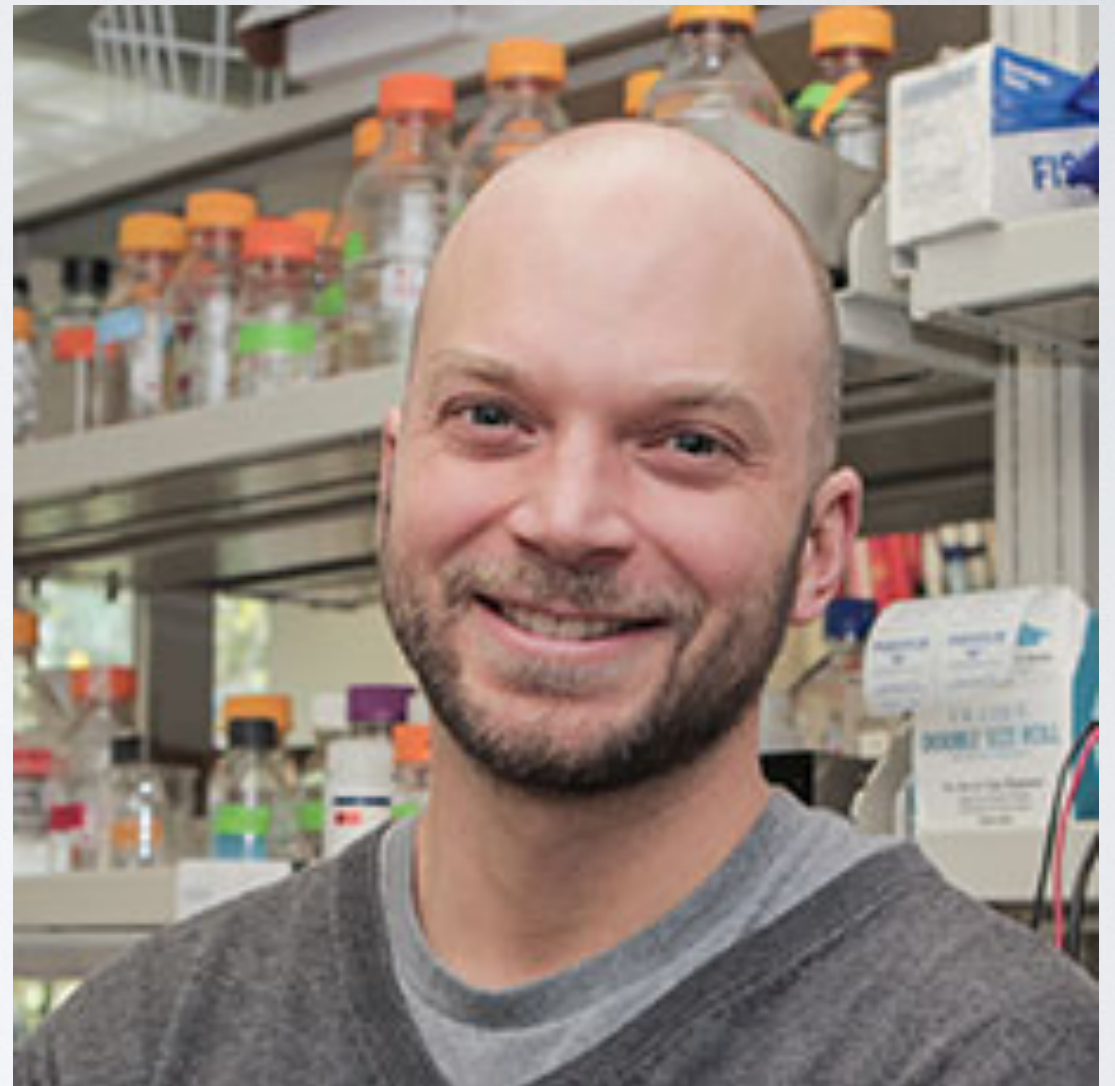


Lecture 3- Vignette 1

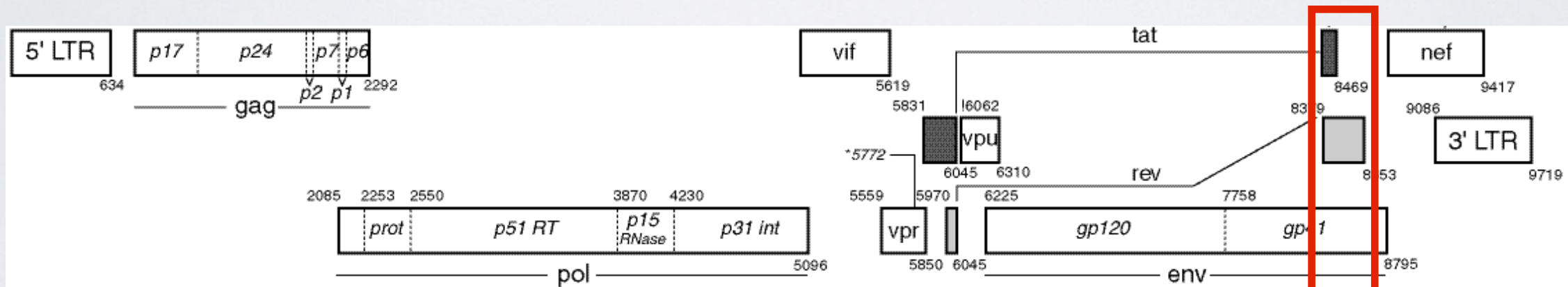
Discovery of a novel overprinting gene in polyomaviruses



Matt Daugherty, former postdoc
now Assistant Professor, UC San Diego

WHAT IS OVERPRINTING? HOW DOES IT EVOLVE?

<i>DNA</i>	CCTATTAATGGAACAGCCA
<i>Frame 1</i>	P I N G T A
<i>Frame 2</i>	L Y M E Q P



hiv.lanl.gov

UNUSUAL EVOLUTIONARY CONSERVATION

Virus A **CCTATTAATGGAACAGCCA**

Frame 1 **P I N G T A**

Frame 2 **L L M E Q P**

Virus B **CCTAT**A**AATGGAACAGCCA**

Frame 1 **P I N G T A** *Synonymous change*

Frame 2 **L ***** M E Q P** *Nonsynonymous change*

Observe depression of ‘synonymous’ mutations in overprinting regions

A NOVEL GENE IN MCPyV

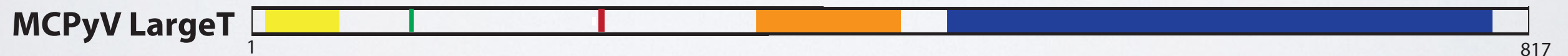
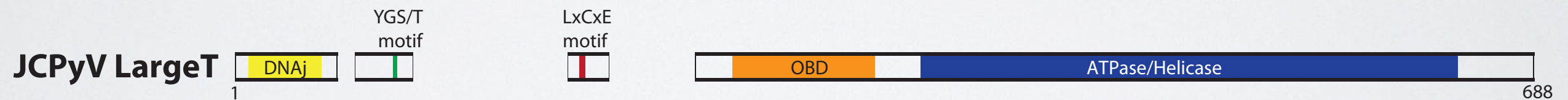
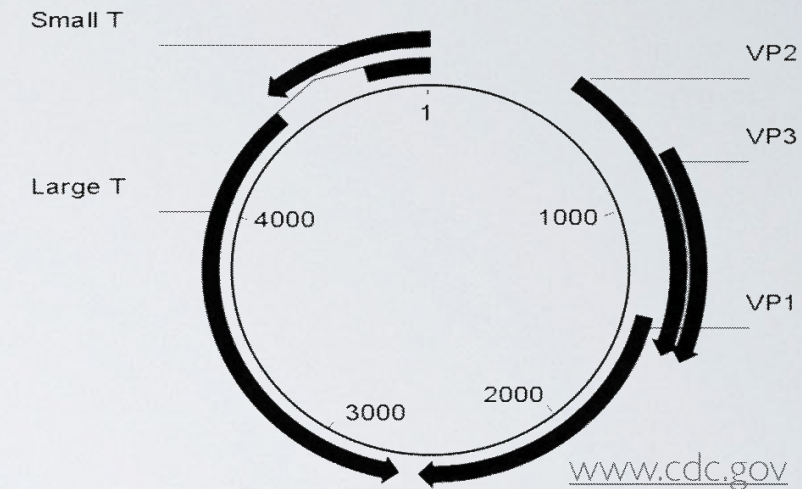
Polyomaviruses (PyV's) are small DNA viruses

9 distinct human viruses, infect >50% of people

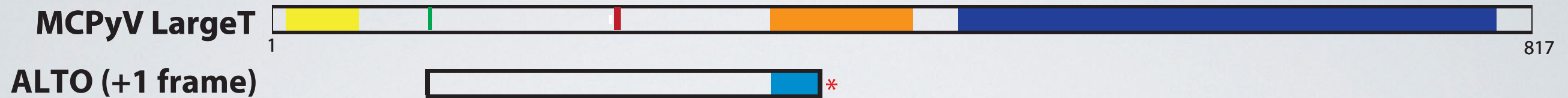
SV40: Causes cancer in monkeys (also has a very useful promoter)

MousePyV: Causes cancer in rodents

Merkel cell polyomavirus (MCPyV): Causes cancer in humans



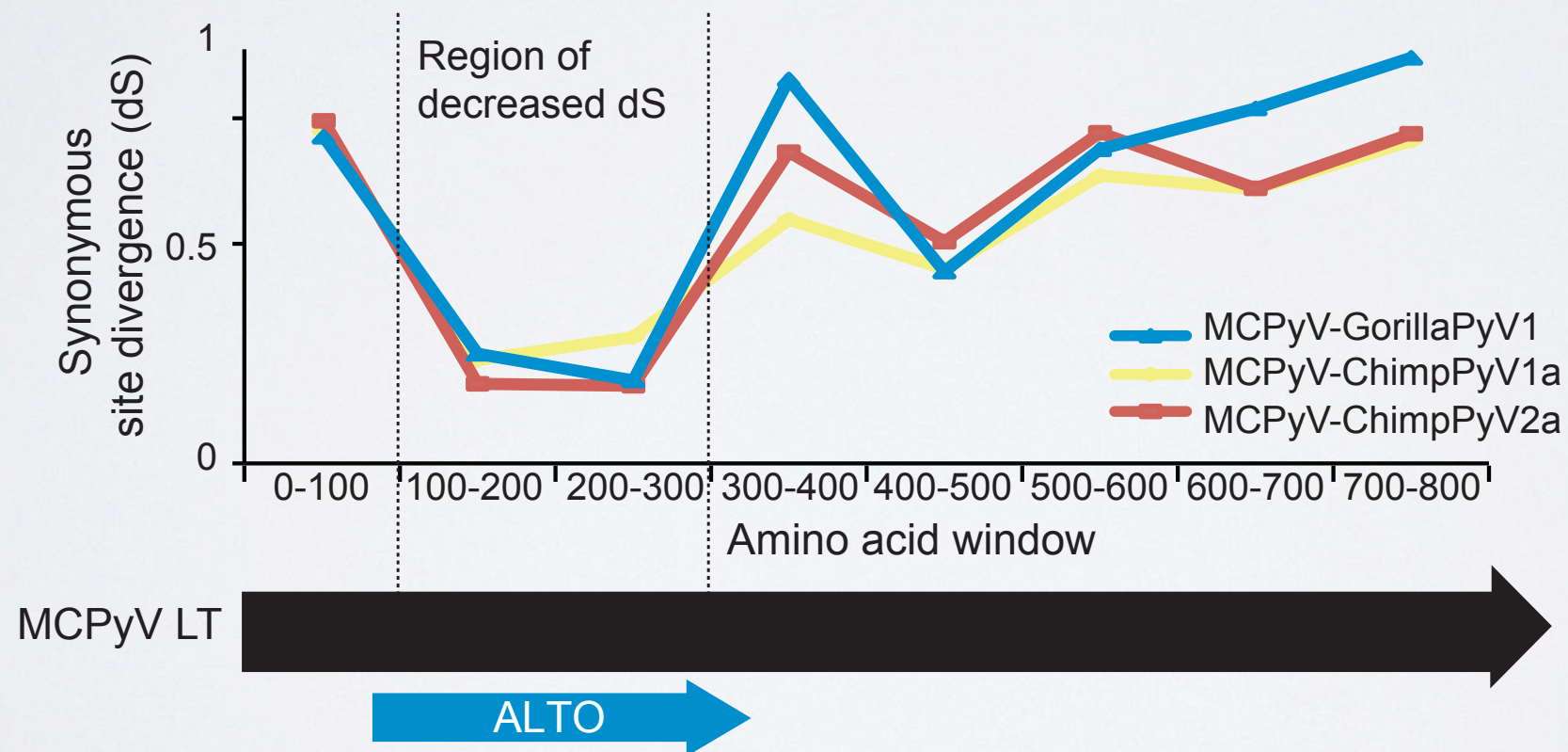
ALTERNATE I ANTIGEN ORF (ALTO)



248 amino acid protein

No recognizable primary or secondary structure except a hydrophobic C-terminus

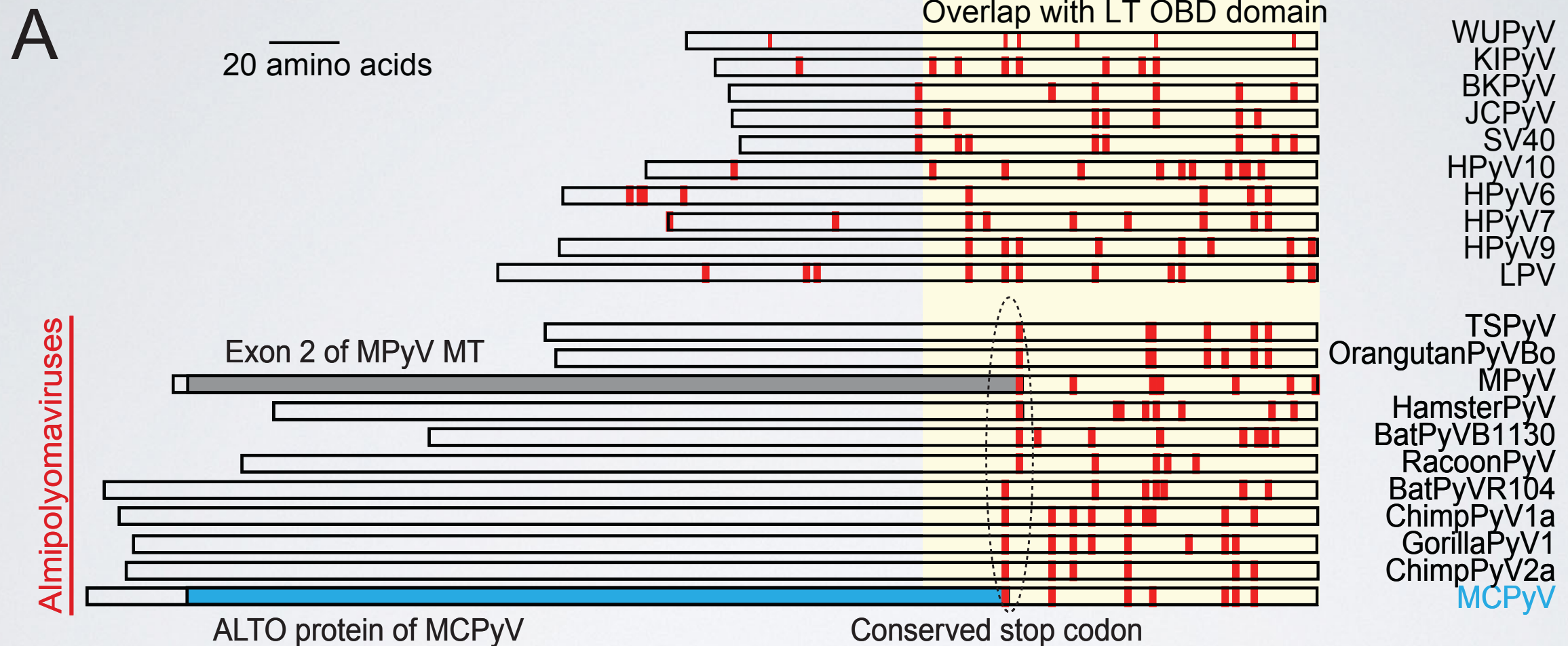
No similarity to any other viral protein except 3 closely related viruses



How did ALTO come to exist?

A WHOLE CLADE OF OVERPRINTED LT GENES

Are there other PyVs with a large uninterrupted ORF overprinting on LT?

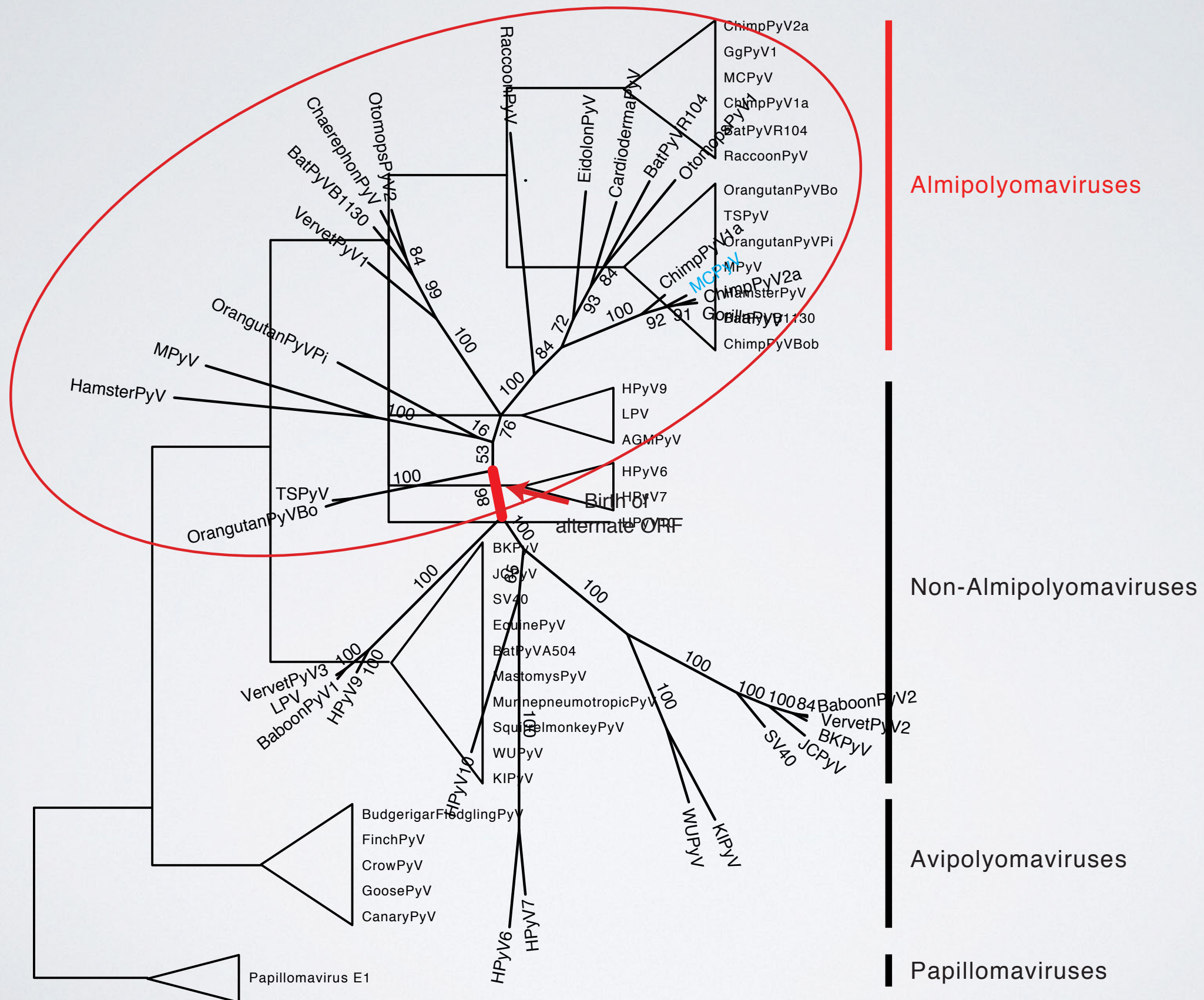


Includes another human virus (TSPyV) and oncogenic MousePyV

The overprinting MousePyV gene (called MiddleT) is sufficient to cause cancer

Defines a clade of Almi(ALTO or MiddleT containing)polyomaviruses

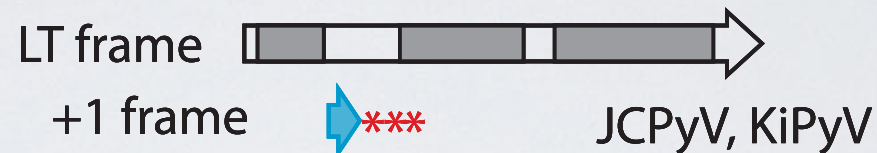
ALTO WAS BORN IN ONE VIRAL LINEAGE



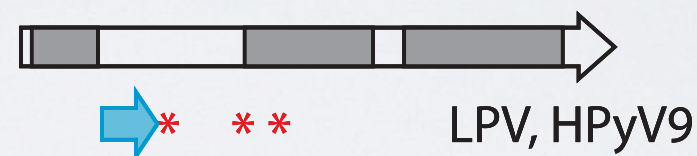
A STEPWISE MODEL FOR OVERPRINTING BIRTH

Ancestral LT gene

*Potential for +1 frame
start codon and
hydrophobic motif
already present*



Expansion of LT protein



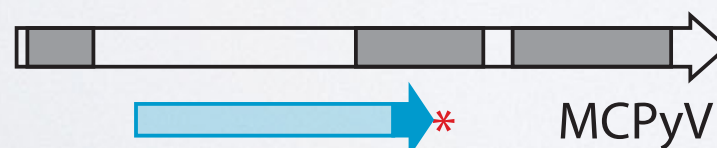
Innovation of ALTO/MT

Removal of stop codons

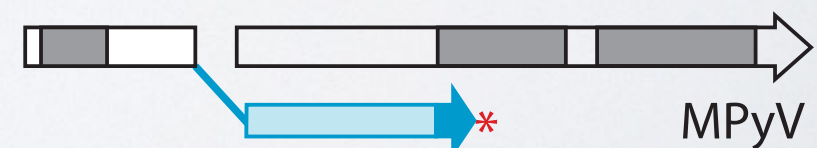


Subfunctionalization of ALTO/MT

Expand LT/ALTO

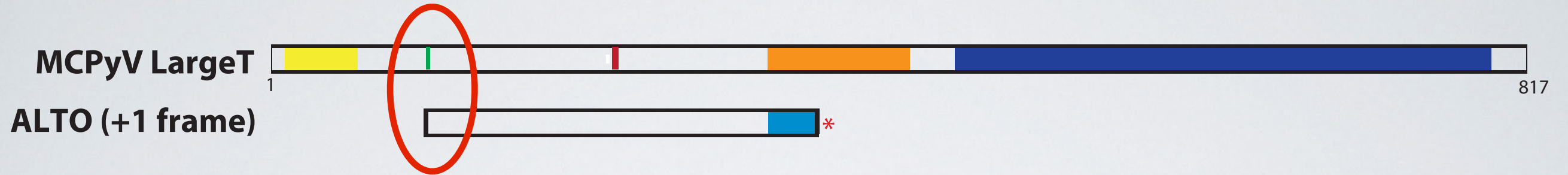


*Create MT through splicing
Expand LT/MT*



HOW WAS THIS NEW GENE BORN?

What was in place before ALTO was born?



LT YGS/T motif

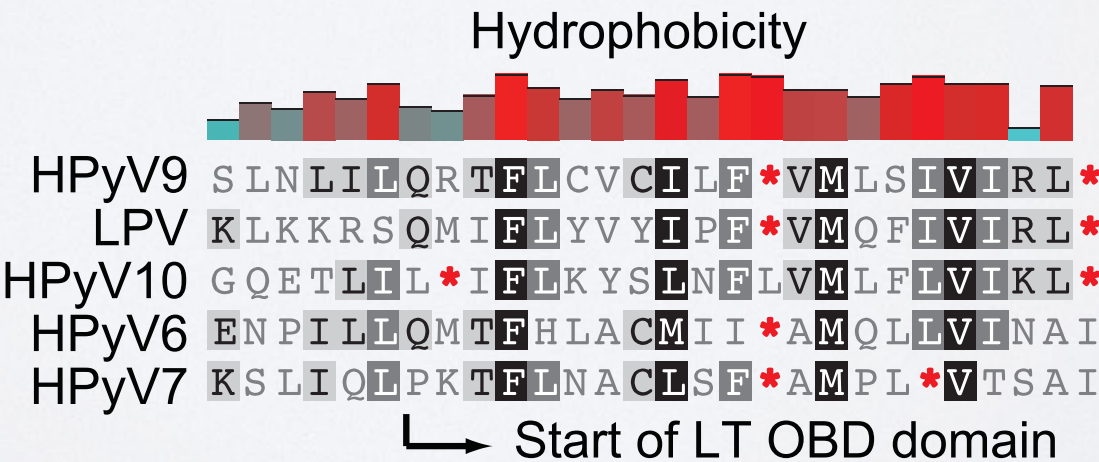
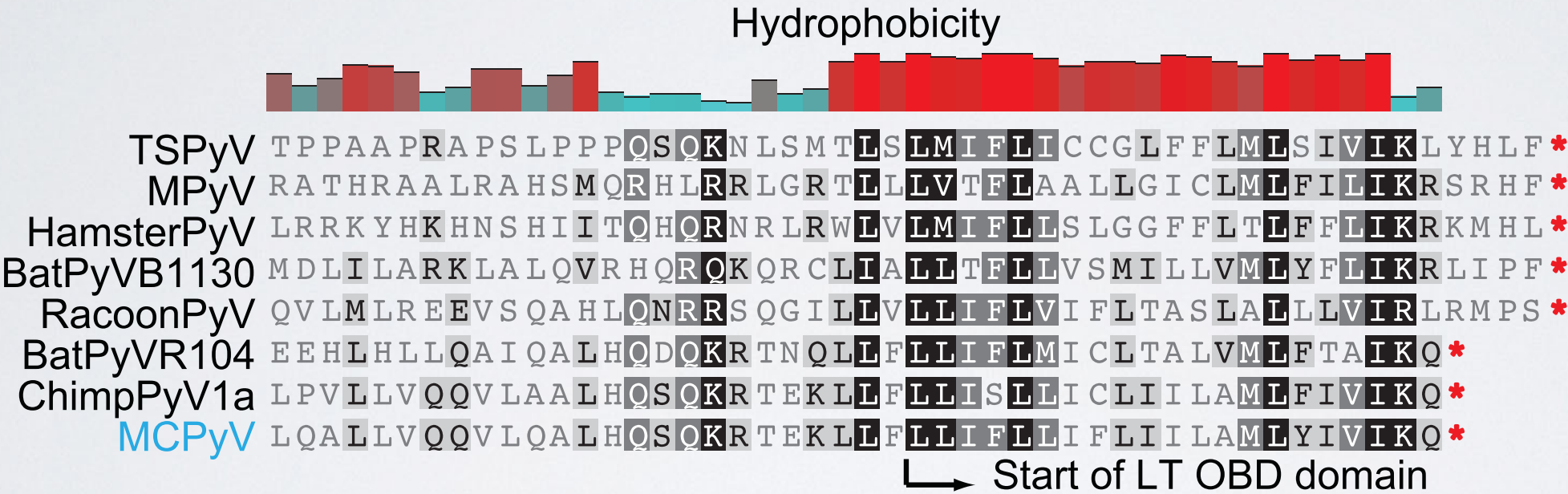
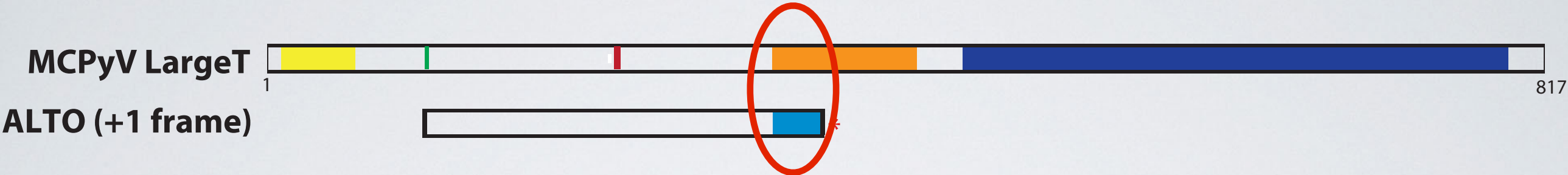
	DNA	LT frame	+1 frame
WUPyV	ATACCCACATATGGGTACCCACAGACTGGGATTACTGGTGGTCTCAGTTTA	I P T Y G T P D W D Y W W S Q F	Y P H M V P Q T G I T G G L S L
JCPyV	GTGCCAACCTATGGGAACAGATGAATGGGAATCCTGGTGGGAATACATTTA	V P T Y G T D E W E S W W N T F	C Q P M E Q M N G N P G G I H L
SV40	ATTCGAACCTATGGGAAGTGTGATGAATGGGAGCAGTGGTGGGAATGCCTTTA	I P T Y G T D E W E Q W W N A F	F Q P M E L M N G S S G G M P L
HPyV6	CCTCCACAATATGGGCAGTCCCGGCTGGGAACAAATGGTGGGCTGATTTCA	P P Q Y G S P G W E Q W W A D F	L H N M A V P A G N N G G L I S
LPV	ATATAGCCCTATGGGATCGAGATGCAGTTCGGGAGGATCTTCCAAATCCA	Y S P M D R D A V R E D L P N P	N I A L W I E M Q F G R I F Q I
TSPyV	ATACCTCCCTATGGGCACCCCTCCTGGGCCAGCTGGTGGGAAAGCTTCA	I P P Y G H P S W A S W W E S F	Y L P M G T P P G P A G G K A S
MPyV	TCATTTCAGCTATGGGAAGCAAGTACTTCACAAGGGAATGGGAATGATTTTC	S F S Y G S K Y F T R E W N D F	H S A M E A S T S Q G N G M I S
RacoonPyV	CCCCCTATATATATGGGAACCCCCGCTTTAGAGAGTGGTGGTTCTACCAG	P P I Y G T P A F R E W W F Y Q	P L Y M E P P P L E S G G S T S
BatPyVR104	GAGCCAATATATATGGTACTACAGCTTTTAGACACTGGTGGTATCGGCAG	E P I Y G T T A F R H W W Y R Q	S Q Y M V L Q L L D T G G I G S
GorillaPyV1	GCCCCCTATATATATGGGAACAGCAAAAATTTAAAGAGTGGTGGCACTCCGGA	A P I Y G T A K F K E W W H S G	P L Y M G Q Q N L K S G G T P E
MCPyV	GCCCCCTATATATATGGGAACCACTAAATTCAAAGAAATGGTGGAGATCAGGA	A P I Y G T T K F K E W W R S G	P L Y M G P L N S K N G G D Q E



ALTO start codon

HOW WAS THIS NEW GENE BORN?

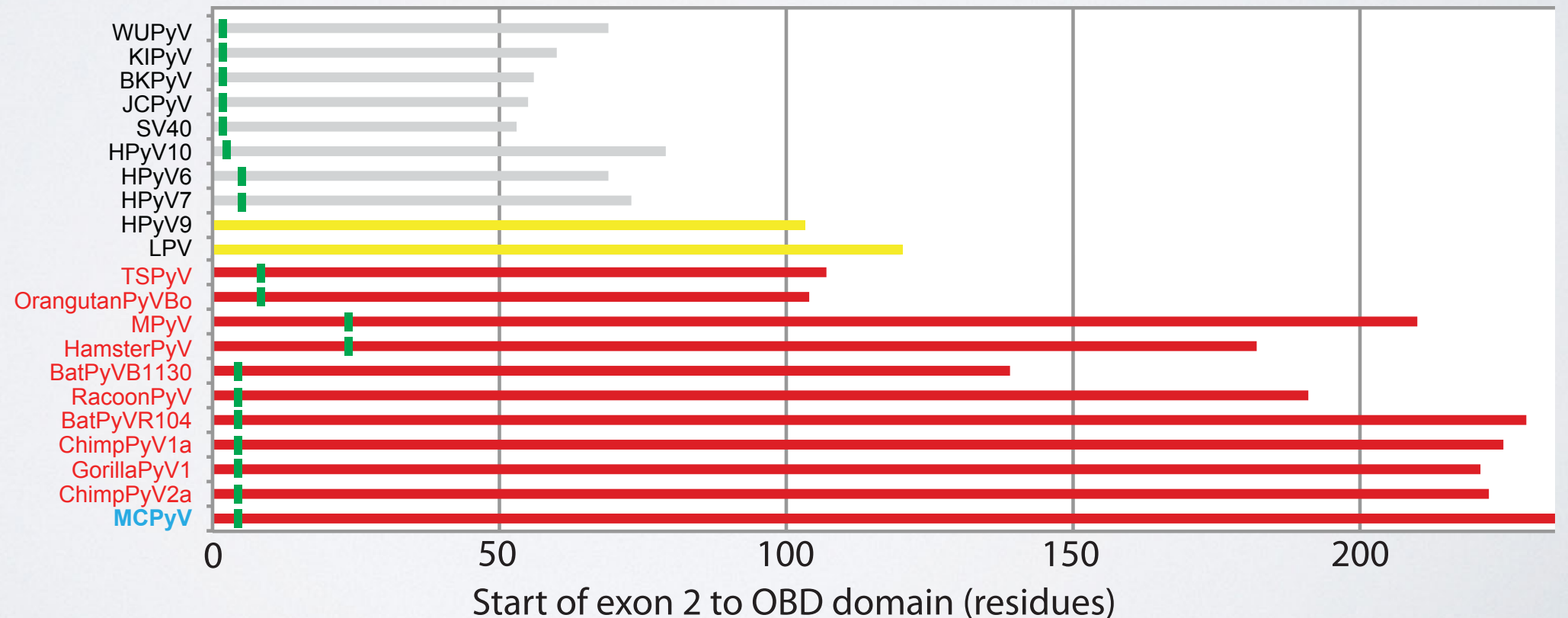
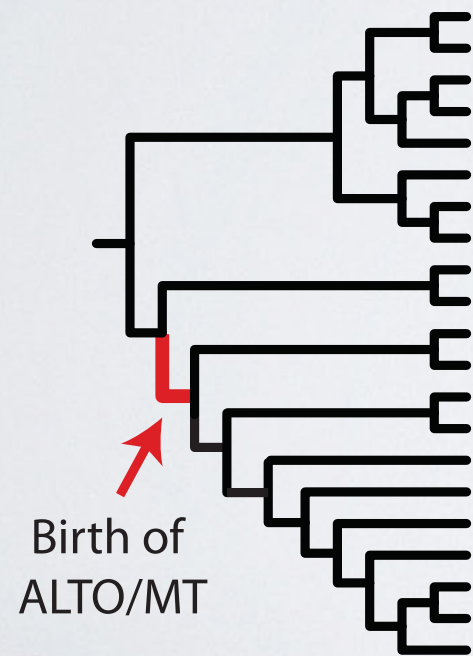
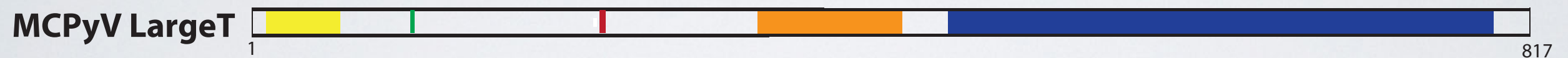
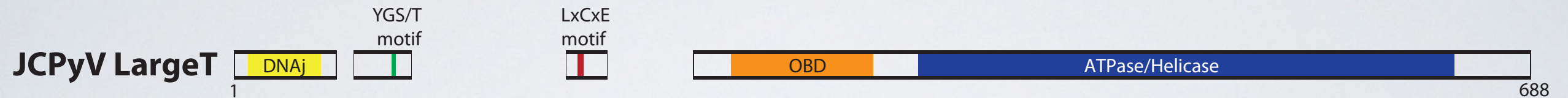
What was in place before ALTO was born?



HOW WAS THIS NEW GENE BORN?

The start codon and hydrophobic motif were in place before ALTO was born

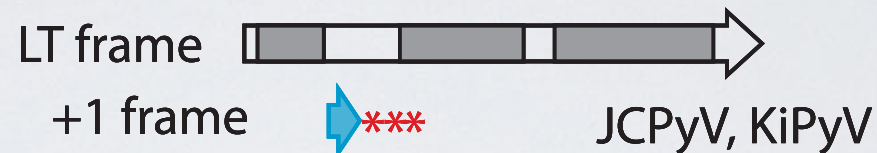
Where did all of the extra sequence come from?



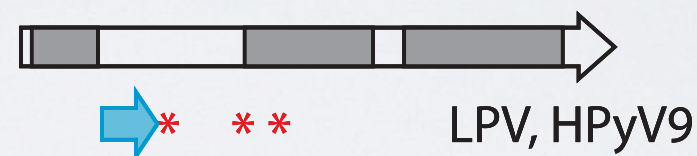
A STEPWISE MODEL FOR OVERPRINTING BIRTH

Ancestral LT gene

*Potential for +1 frame
start codon and
hydrophobic motif
already present*



Expansion of LT protein



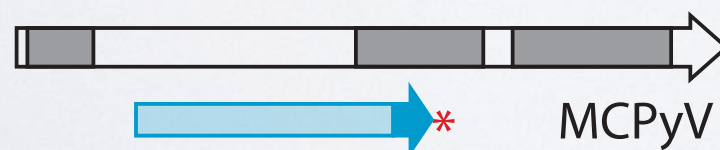
Innovation of ALTO/MT

Removal of stop codons

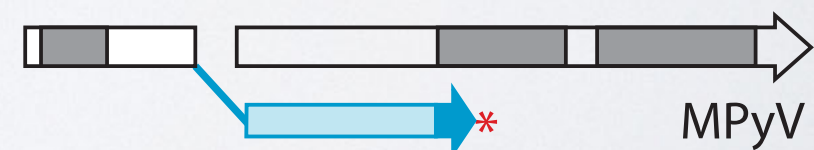


Subfunctionalization of ALTO/MT

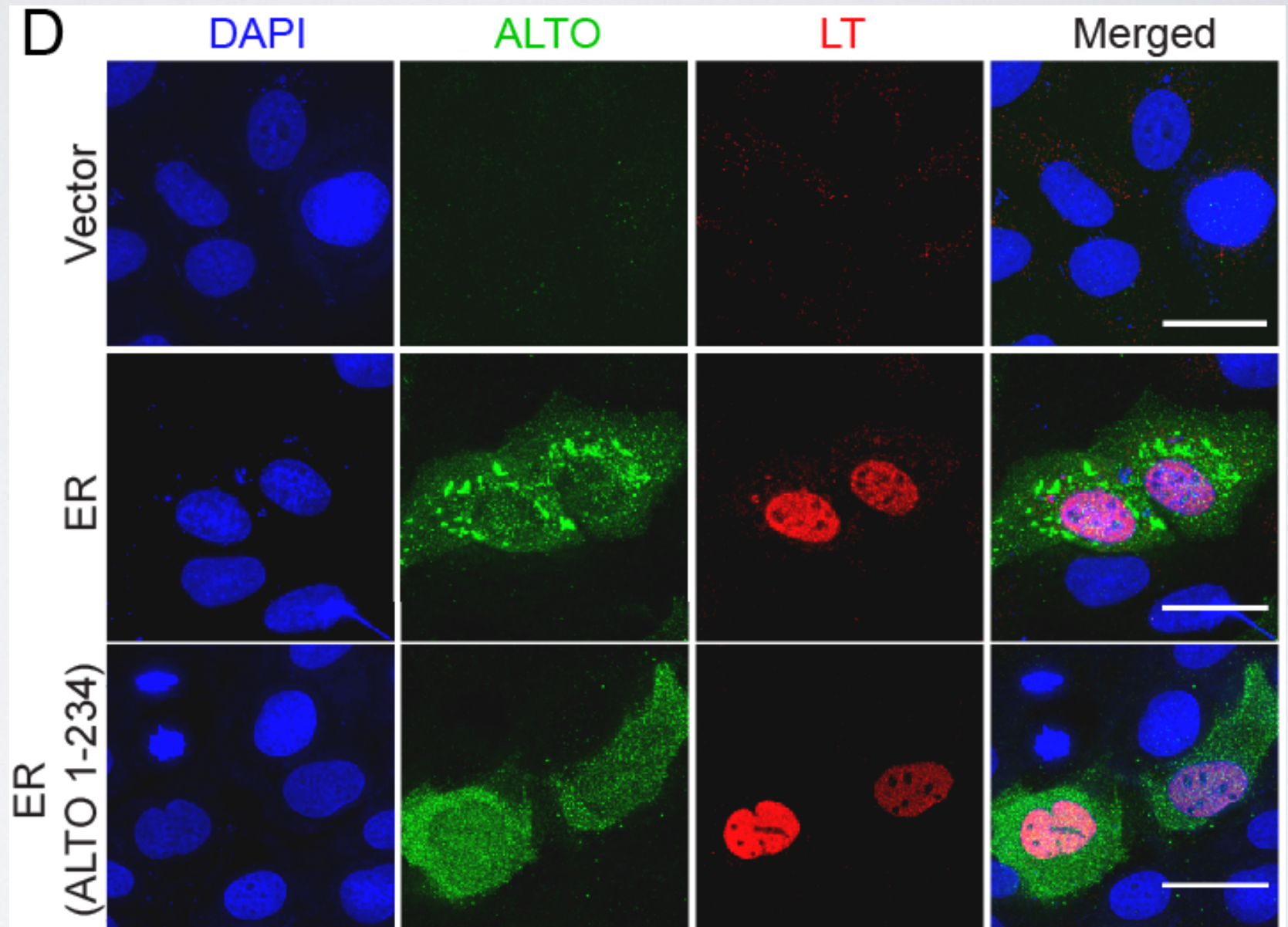
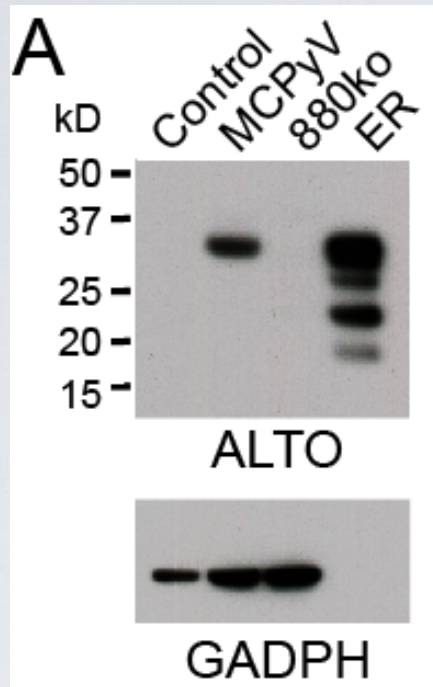
Expand LT/ALTO



*Create MT through splicing
Expand LT/MT*

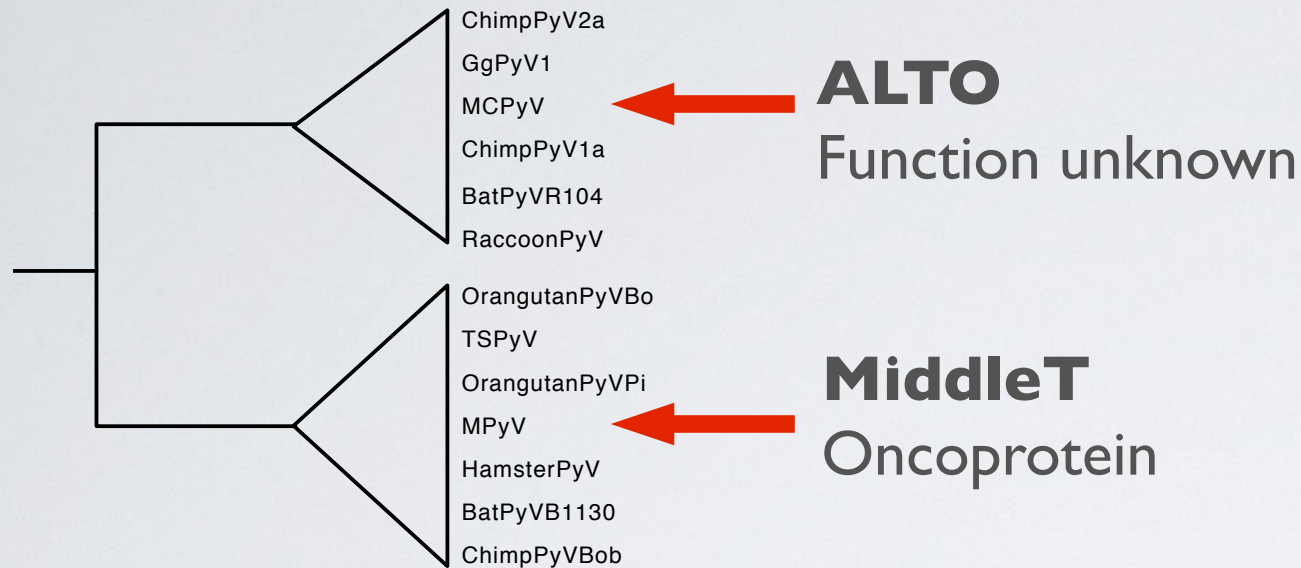


WHAT DOES ALTO DO?



The short answer: we don't know!

WHAT DOES ALTO DO?



Evolutionary connection to MousePyV allows us to use MiddleT to understand ALTO

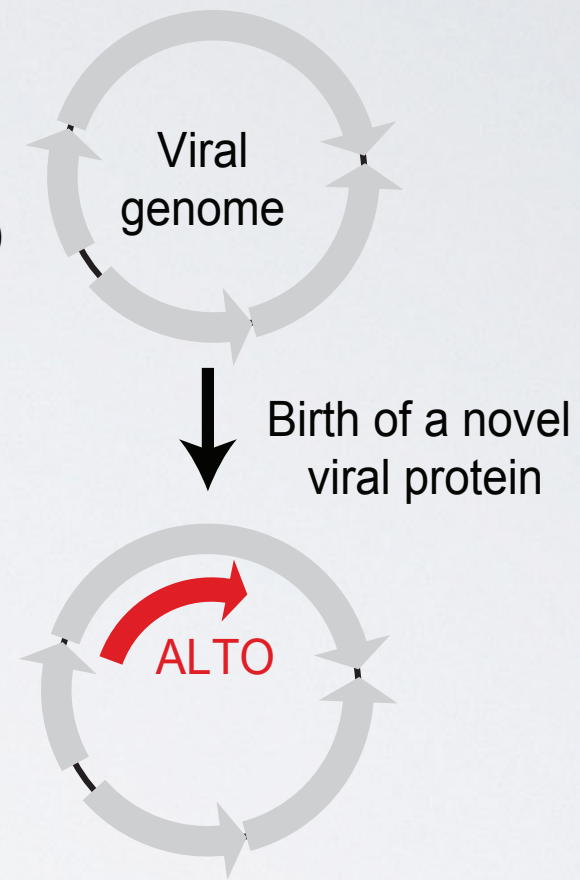
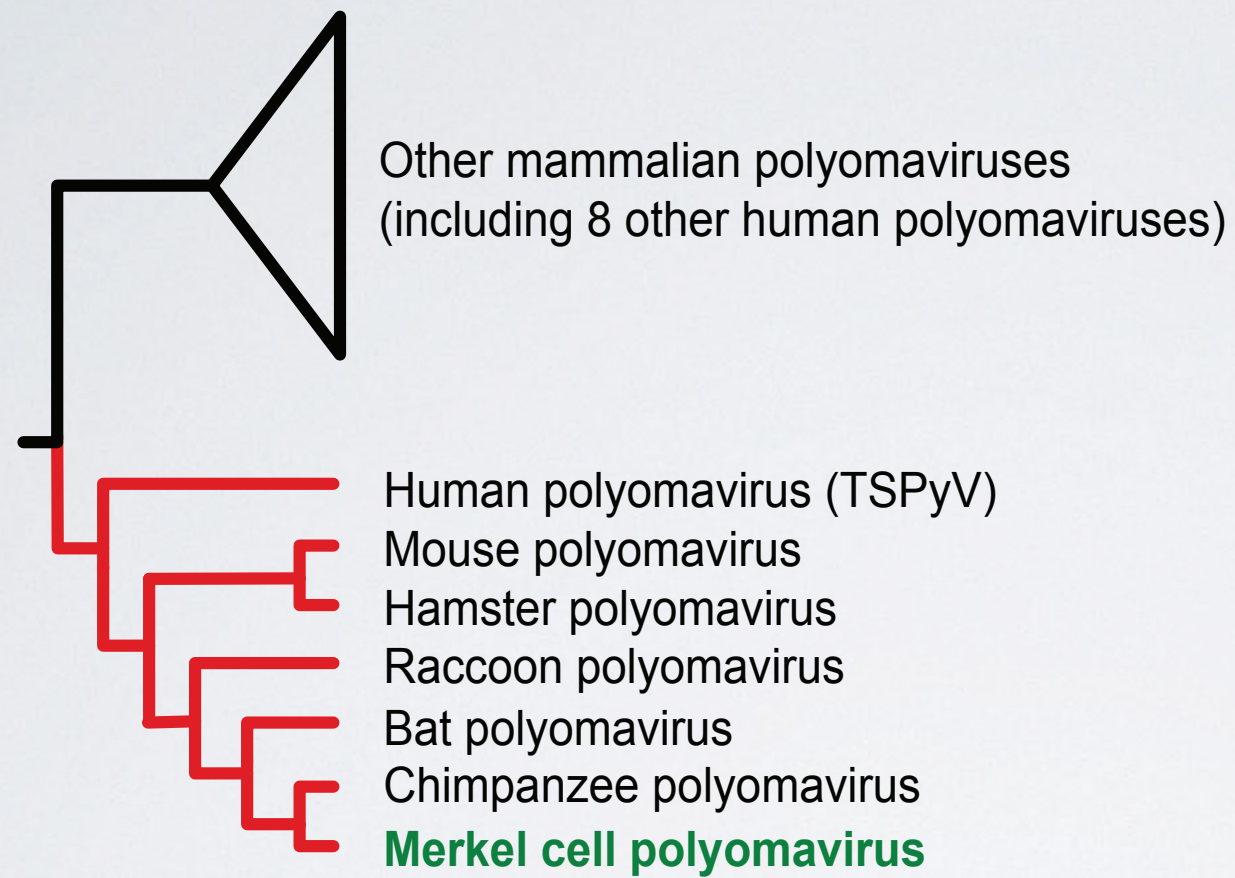
MiddleT transforms cells using short linear motifs to disrupt cell cycle control

Several of these motifs have phosphorylated Y's

Does ALTO have conserved Y motifs?

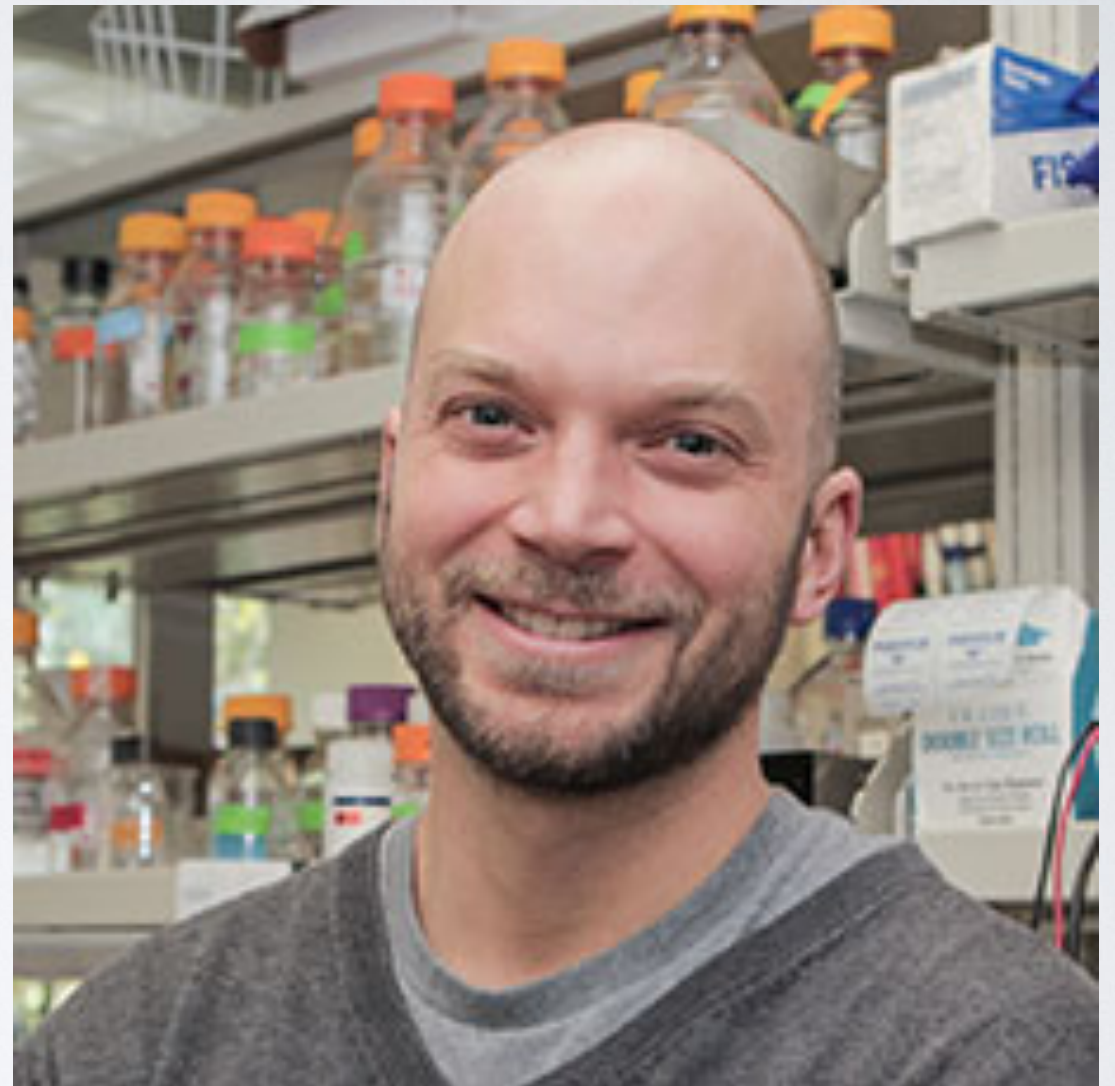
Do these motifs activate signaling cascades?

RaccoonPyV	-----MNPLMEPPPLESGGSTSRITYTPGVMMAR-----PPQVPQEDPE---D
BatPyVR104	-----MKSQMVLCQLLDGTGGIGS--TAEAMAGPR-----LPLQEQAPDPQVMVP
OtomopsPyV1	MMRRMMMRIPREMALOPSERGEGGS-----NMAIGE-----QRFPQVPEE-----E
EidolonPyV	---MKRRMIPLEMCPPNSRPGGTRS-----TPAGARGAMTLPHLNHPVEDPQDVPDRTVG
CardiodermaPyV	-MILKKKTSPEHMAPHNLDIGGLSS-----STMGLN-----LQLMQGPVEIGEE
ChimpPyV1a	-----LTSLLMAQQNLSSGSLPE-----GMALGR-----HTPVEDPESPRQTHPG
MCPyV	-----LTRPLMCPLNSKNGGDQE-----DSASGRHMNMGPHTGTQDPESLPPMHFG
GorillaPyV	-----STRPLMGQQNLKSGGTPE-----AMALGR-----RTRPVQDPESPPPMHFG
ChimpPyV2a	-----STRLLMEQFDSKSGGIQD-----SMALGR-----HTPVQDPESPPPLHFG
	* * . 1
RaccoonPyV	GIPLGIMENSPEARGPAGGT---GYIPQS-----QVF-----
BatPyVR104	N-PANHHPLGARMDSGSSGSMESQETQTTLLTNSSITNTSLTIVNSSTSRTPSRSSSS
OtomopsPyV1	SLPIQETPAAMAVVNSPRGP---QNFQSH-----PIF---ET
EidolonPyV	TRP-----RSRTPTPPL---FHRPRL-----PTPQMDSPS
CardiodermaPyV	SLPLPTQEQGAPRGPLPDI---LQRPRL-----PTP---PQT
ChimpPyV1a	GPPVT-EPPLPAARALPLDMQQAHSRPRRL-----QTP-----S
MCPyV	EPPVEAH--HPTARALPLGM---GPSQRPL-----QTP-----S
GorillaPyV	EPPVE-PPLQPAARALPPAM---GPSQRPL-----QTP-----S
ChimpPyV2a	EPPLTEHPLQAAARALPLAM---GPSQRPL-----QTP-----S
	* *
RaccoonPyV	-----LCLPTD-----P-----
BatPyVR104	-----PRNPITHPLPLAPD-----QTYLPHMARM---
OtomopsPyV1	-----PPERIYLPHLL-----P-----
EidolonPyV	SGSLPSIPEEPLYLPRDFNPTRDTIMDLDRSHQVDGGRRSAP---PAMLPRENS---
CardiodermaPyV	-----PTEPIYLPQPLMMRRQD---PLPR-----PPGLP---FPTLNPPPIQE
ChimpPyV1a	-----PEDPIYLPNSLQNPHPH---LQPR-----P---P---IPEENPA---
MCPyV	-----PEDPIYLPNTMRNPPLDPVAERR---P---P---IQEENPA---
GorillaPyV	-----PEEPVYLPNTLQRLP---MEPR-----P---P---VQEENQA---
ChimpPyV2a	-----PEEPVYLPNTLMTTP---MQQR-----P---P---IQEENPA---
	: *
RaccoonPyV	LxCxE... -TPQPIYTEIIPRTHSSVKRHLINHHHHHHHRLLTPTPSPP---LSPLLPP---RLSD
BatPyVR104	GSQEGIYMEIHPESH---GMIPFY-ARS---PYHRP---LLQSQP-----
OtomopsPyV1	SSREGTYLEIQQDVPLPRPGMTSPSAVMRP-CLHPNPRPPRRSPSLNHQGRG-----
EidolonPyV	GGEEAIYLEIAPEY---PHLIGMHCDA-MSP-C---PRPRS---LQVRRTP---RAAQ
CardiodermaPyV	GTEAAIYLEILQED---PSQIGMI-PYT-ATS-LCPSPRPFRP---MRHFRSPSPRPPSP
ChimpPyV1a	HPAEPVYLEVLPGAVAP---GMI-STA-MSP-SPHQSLLPP---LRSPRSP---PAP
MCPyV	HPMEPVYLEILPERMAP---GRI-SSA-MNHFPPLSLPRP---LRSLRSP---PPQE
GorillaPyV	HSQEPVYLEVLPEVMAP---GMI-CSV-MNFPARQSPFRP---QRNLRSP---PSQQ
ChimpPyV2a	HNTEPVYLEVLPEVMAP---GMI-CSV-TSRFPRLSLPRP---QKSQRSP---PPLQ
	: * *: *
RaccoonPyV	 SPVRHRPR-PMPPREI-----PGPPDPIPTLLRL---ADENTERIIQQVLMRLRE
BatPyVR104	-----PLPRRNPFVLSQVL---FEDPPGPLRSRLTTTSQCQEVLLSEEHHLHL---
OtomopsPyV1	-----PLPARNP-----MR---TLSPPEVHPILI---EGEGLQD-----LLH
EidolonPyV	SPPRNTPRIPLPKRPPSRNRSTSTLDPPALPSQLQEPLLVEESSLDPRDFVLRLMQ
CardiodermaPyV	RPRPREPALPLPRRP---SPLM---NPSCHQAH---HDENGT-----LLQH
ChimpPyV1a	EGRRSSPRIPLPRRPPLHLNLQMR---NAEDLHSPHQRPPLSPSESENSGGTEALPVLLVQ
MCPyV	-ARPGSPRLPLPRRPRLSLQMR---NTDPPSPPRRLLHSQESSENLGGEALQALLVQ
GorillaPyV	GAAPSSPRPLPLKRPRLSLQMR---NVEPPSPQRPLLP---ESENSGGFEVLQALLVQ
ChimpPyV2a	GAAPSSPRPLPLKRPRLSLQMR---NTEPLHSPPRRPLLPSPSESENSGGFEALQALLVQ
	*: *
RaccoonPyV	OBD domain..... EVSQAHLLQNRSSQGILLVLLIPLVIFLTASLALLLVIRLRMP
BatPyVR104	QAIQALHQSQKRTNQLLPLLIPLMICLTALVMLFTAIRK---
OtomopsPyV1	QVILAHHQSQKSLRIILLIPLVILQIVLIILATLSIVIRL---
EidolonPyV	IAIQALRRGQRNQKRMIFLMIPLLIILLITILVMLSTAIR---
CardiodermaPyV	RLTLALLQNRKTKGMLLLMTPLLIISLNLVMLYILIKH---
ChimpPyV1a	QVLAALHQSQKRTKLLPLLIISLLICLIILAMLFIVIKQ---
MCPyV	QVLQALHQSQKRTKLLPLLIPLLIPLIILAMLYIVIKQ---
GorillaPyV	QVLQVLHQSQKRTKLLPLLIPLLIPLIILAMLYIVIKQ---
ChimpPyV2a	QVLQALHQSQKRTKLLPLLIPLLIPLIILAMLYIVIKQ---
	 :*: * *



Lecture 3- Vignette 2

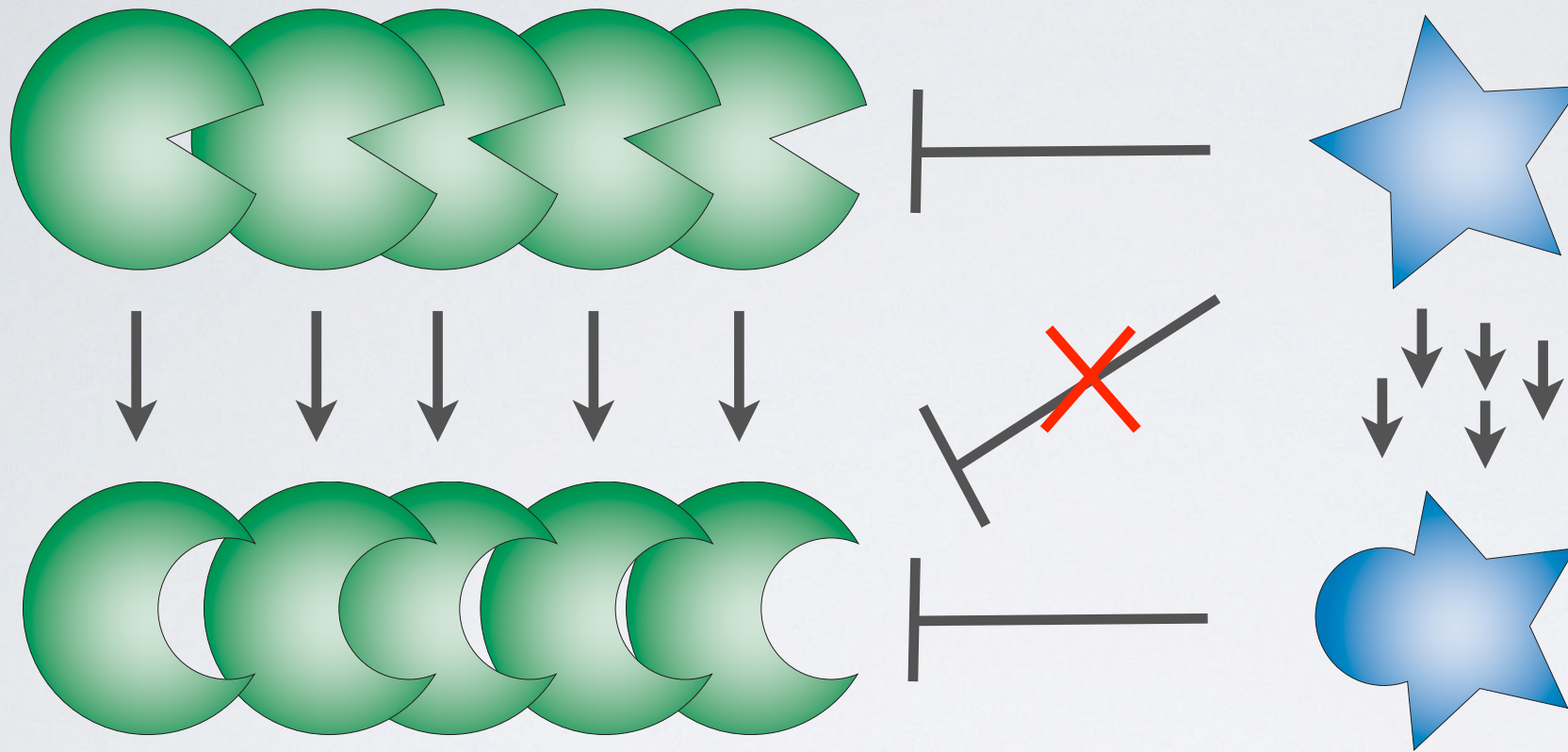
Deconvolution of two ancient mammalian antiviral families



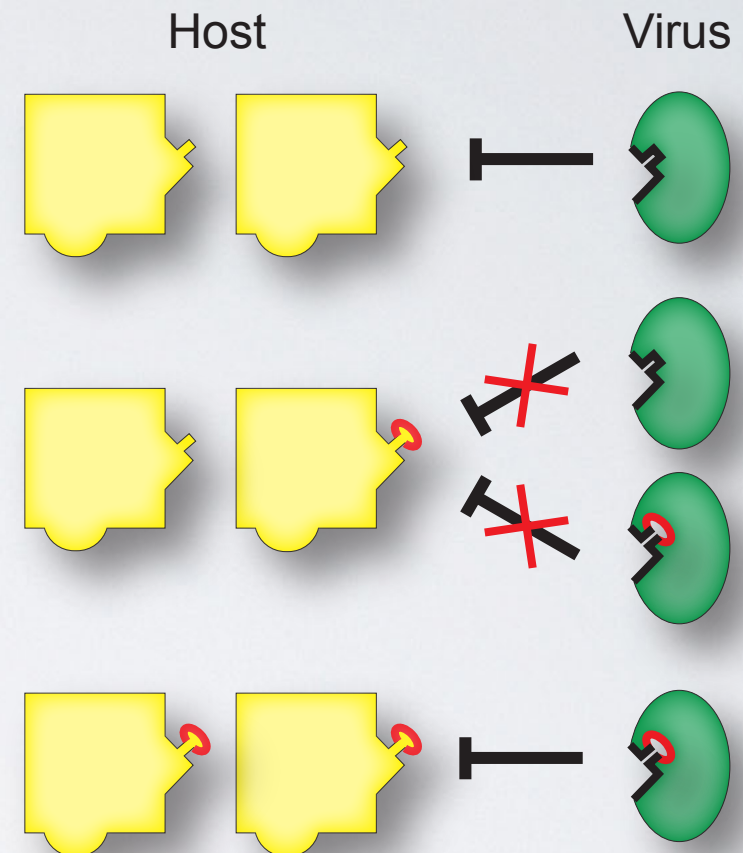
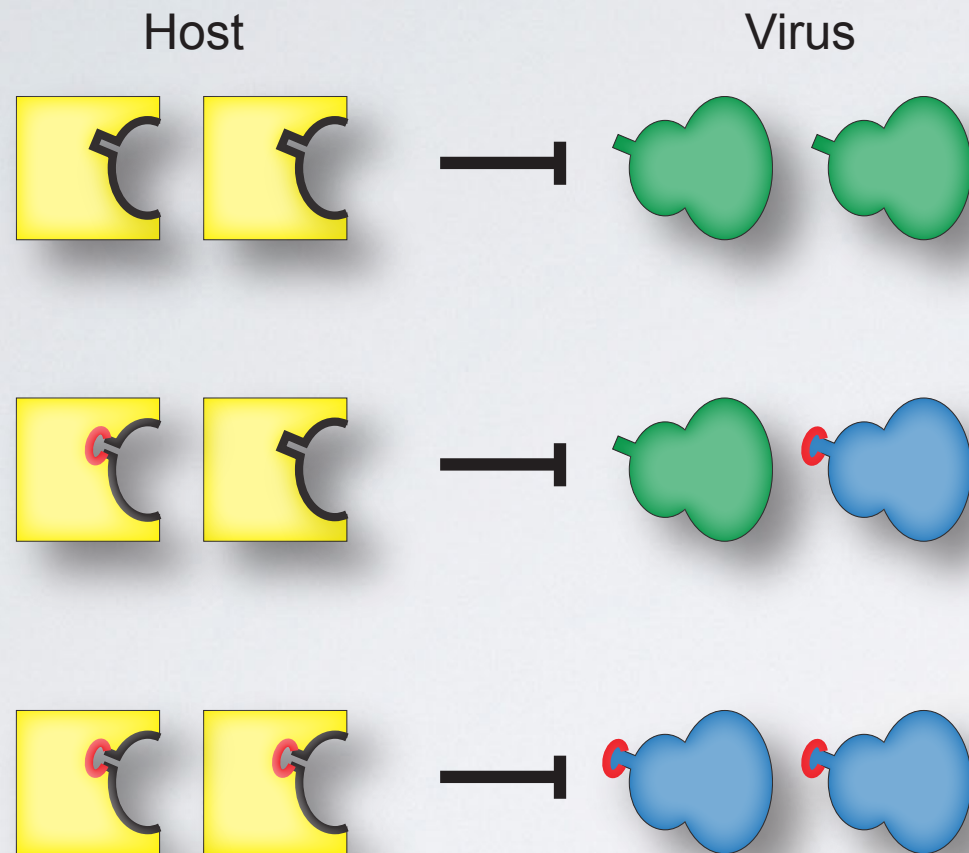
Matt Daugherty, former postdoc
now Assistant Professor, UC San Diego

Antiviral protein

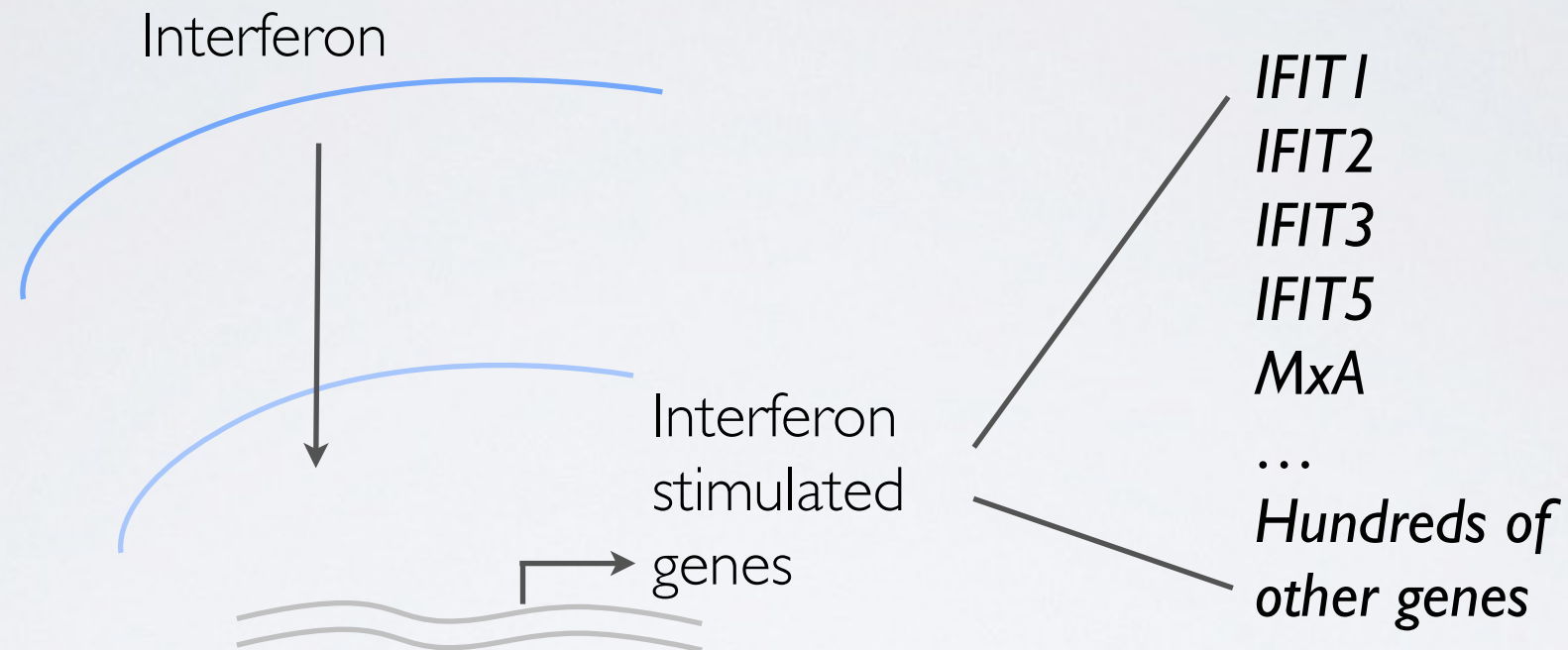
Viral protein



Heterozygosity or gene duplications provides host genomes with an advantage against viruses

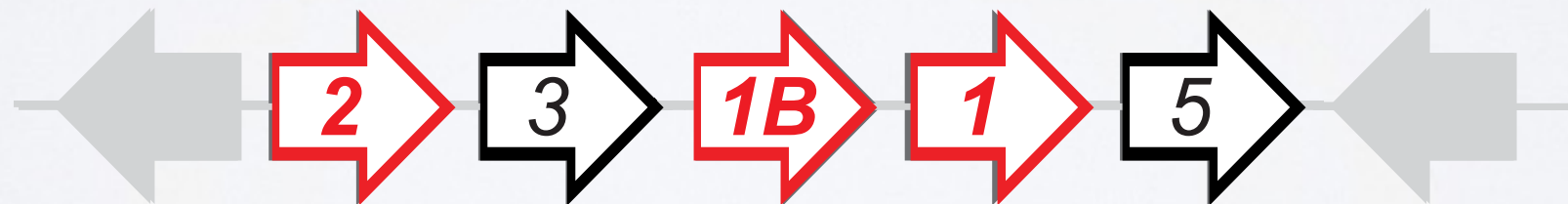


IFIT genes: a case-study in diversity

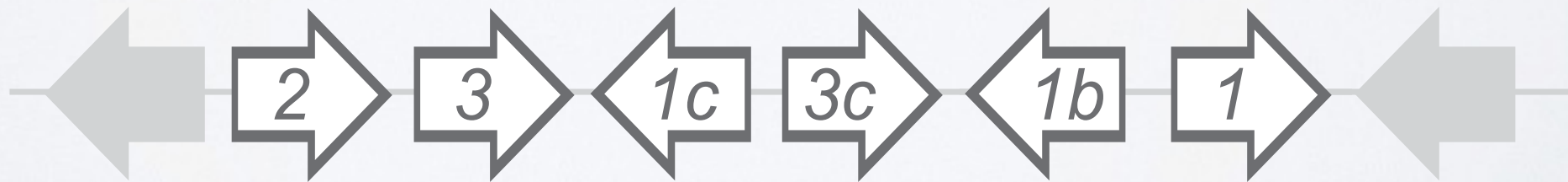


IFIT: Interferon-induced with tetratricopeptide (TPR) repeats

Human
Chromosome 10

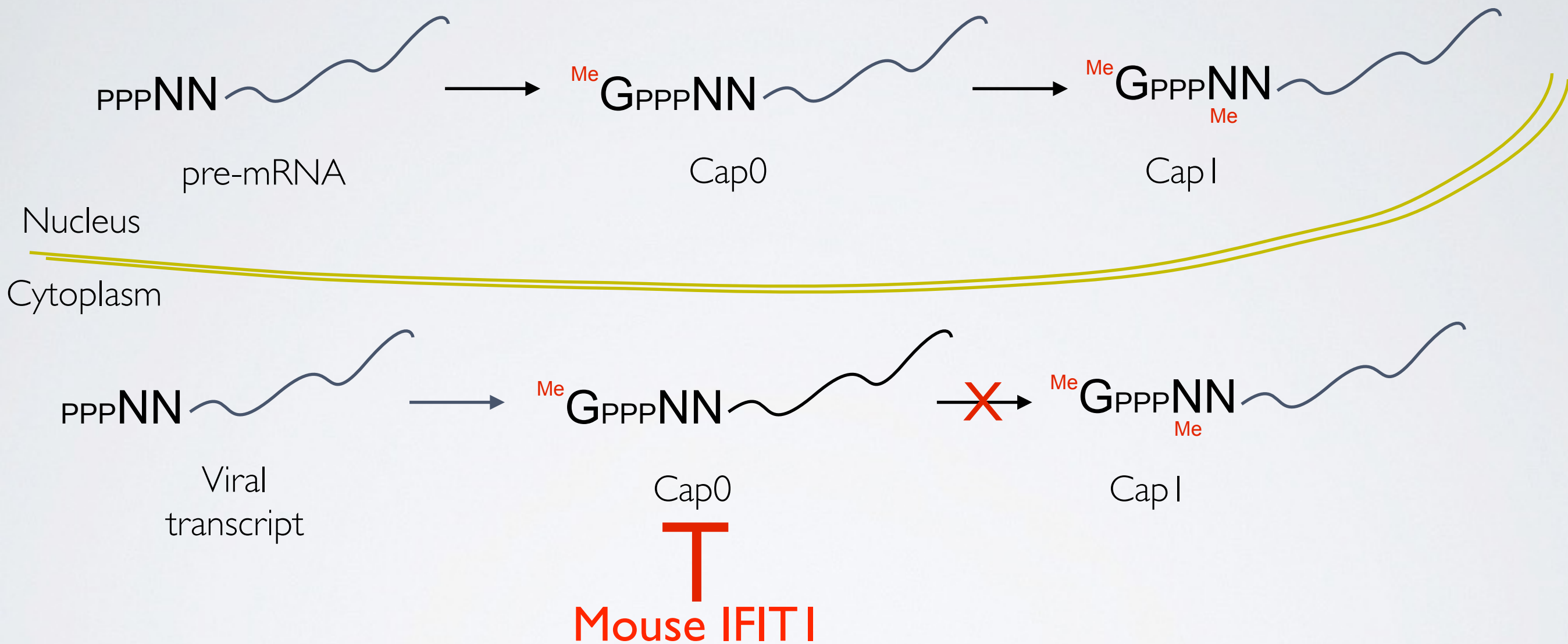


Mouse
Chromosome 19

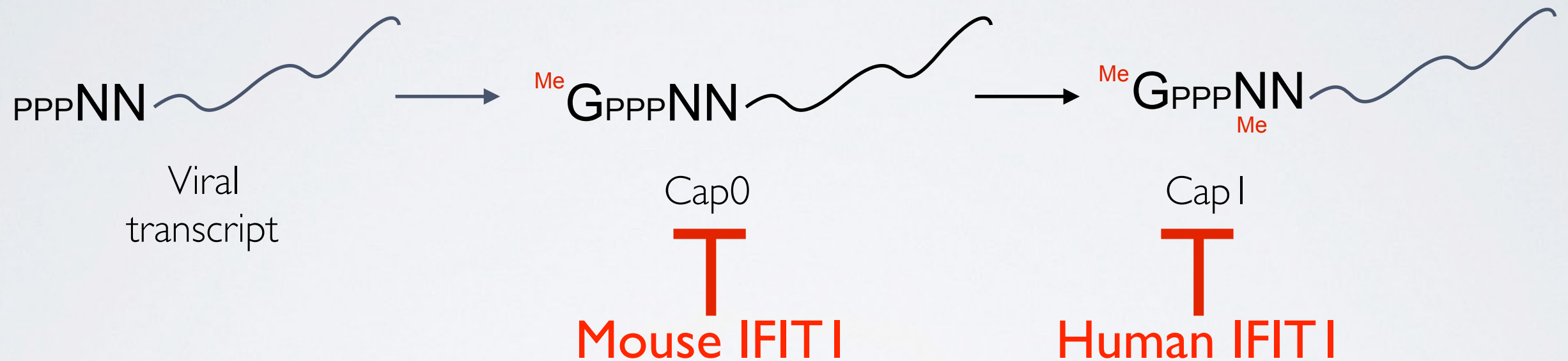


What are the functional consequences of IFIT evolution?

IFIT1 distinguishes 'self' from 'non-self'

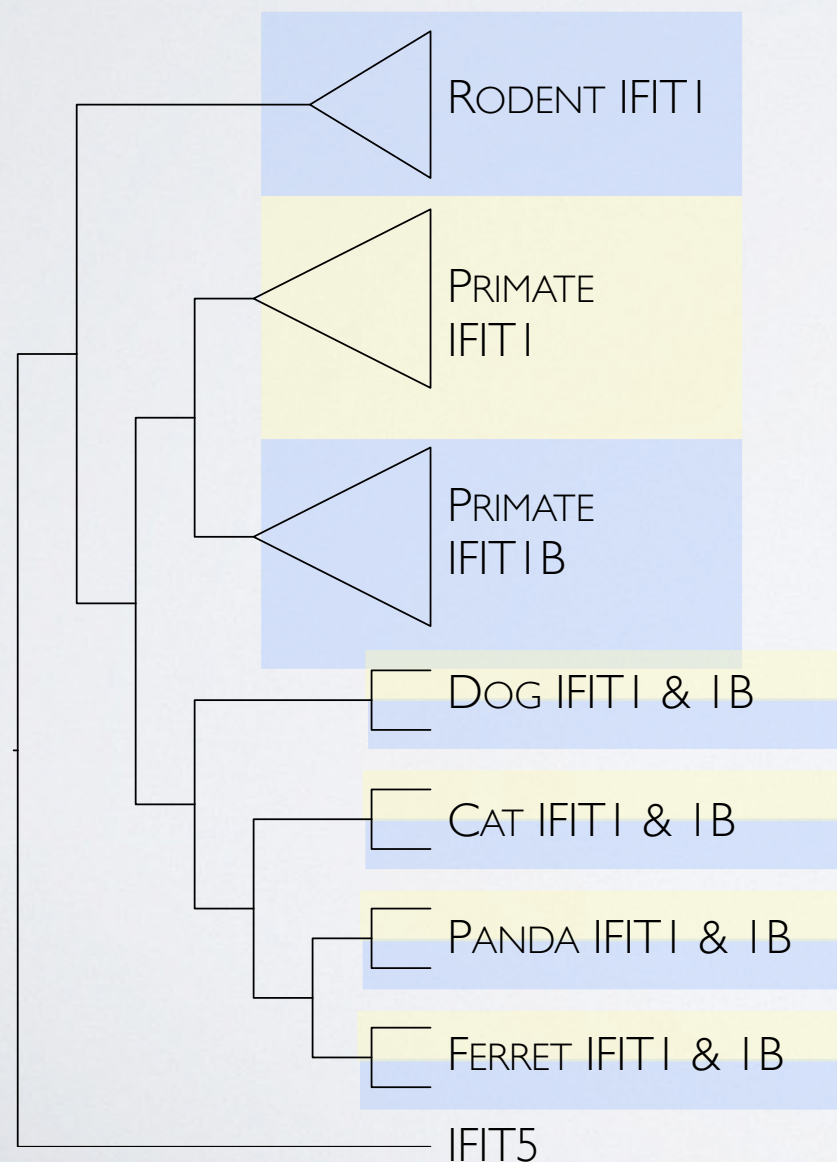
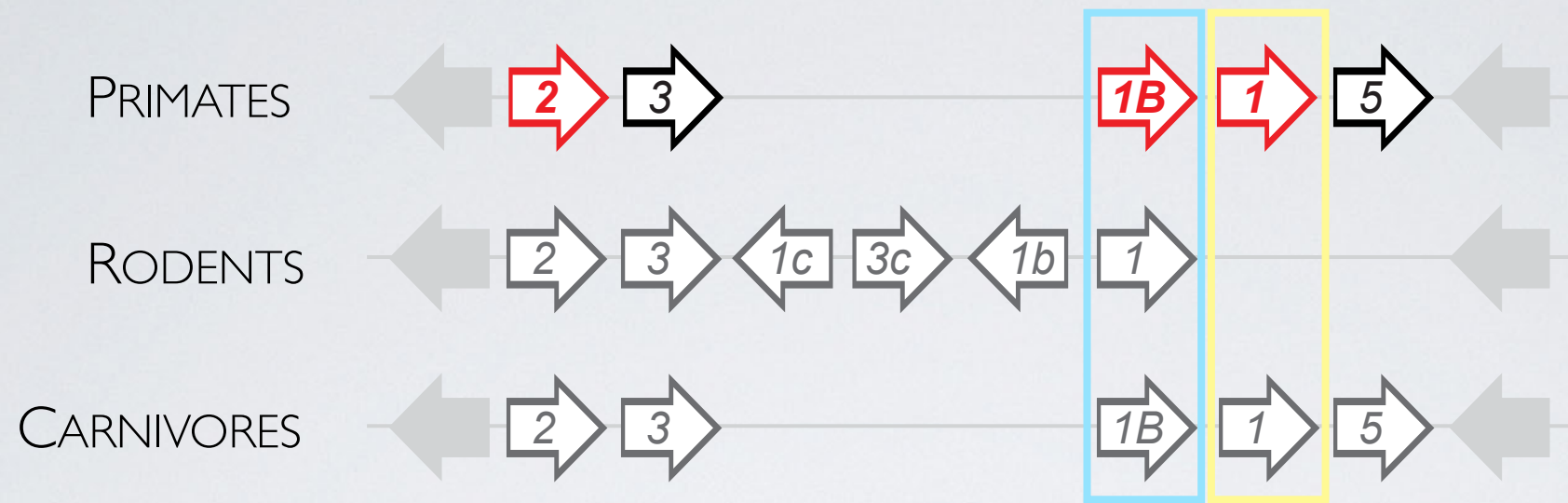


A discrepancy in IFIT1 function



Pichlmair, 2011
Andrejeva, 2013

IFIT phylogeny violates mammalian phylogeny

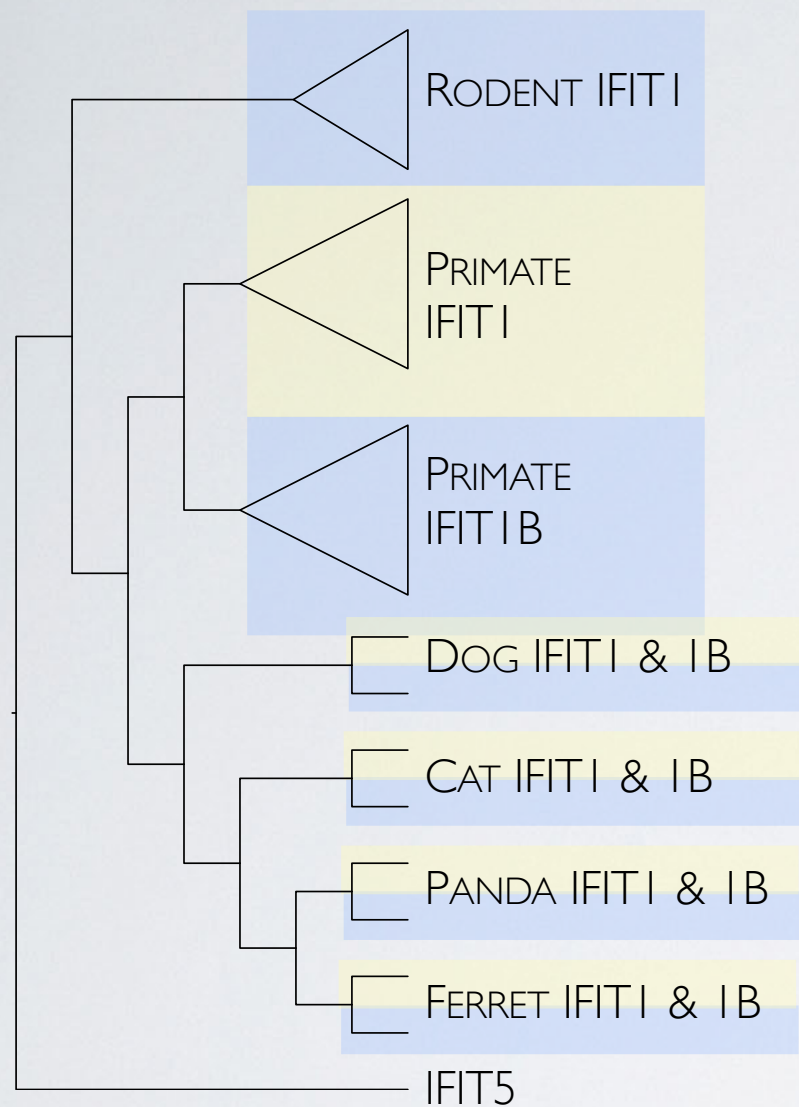


Whole gene phylogeny of IFIT I and IFIT IB

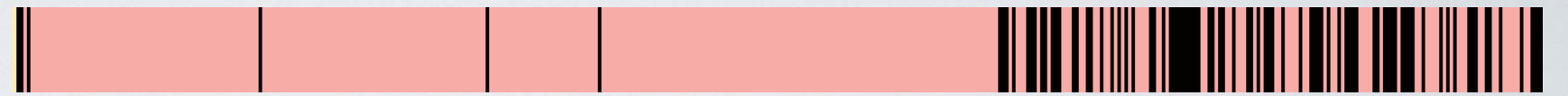
Suggests many independent duplications
(into the same chromosomal location)

Is there another possible explanation for this phylogeny?

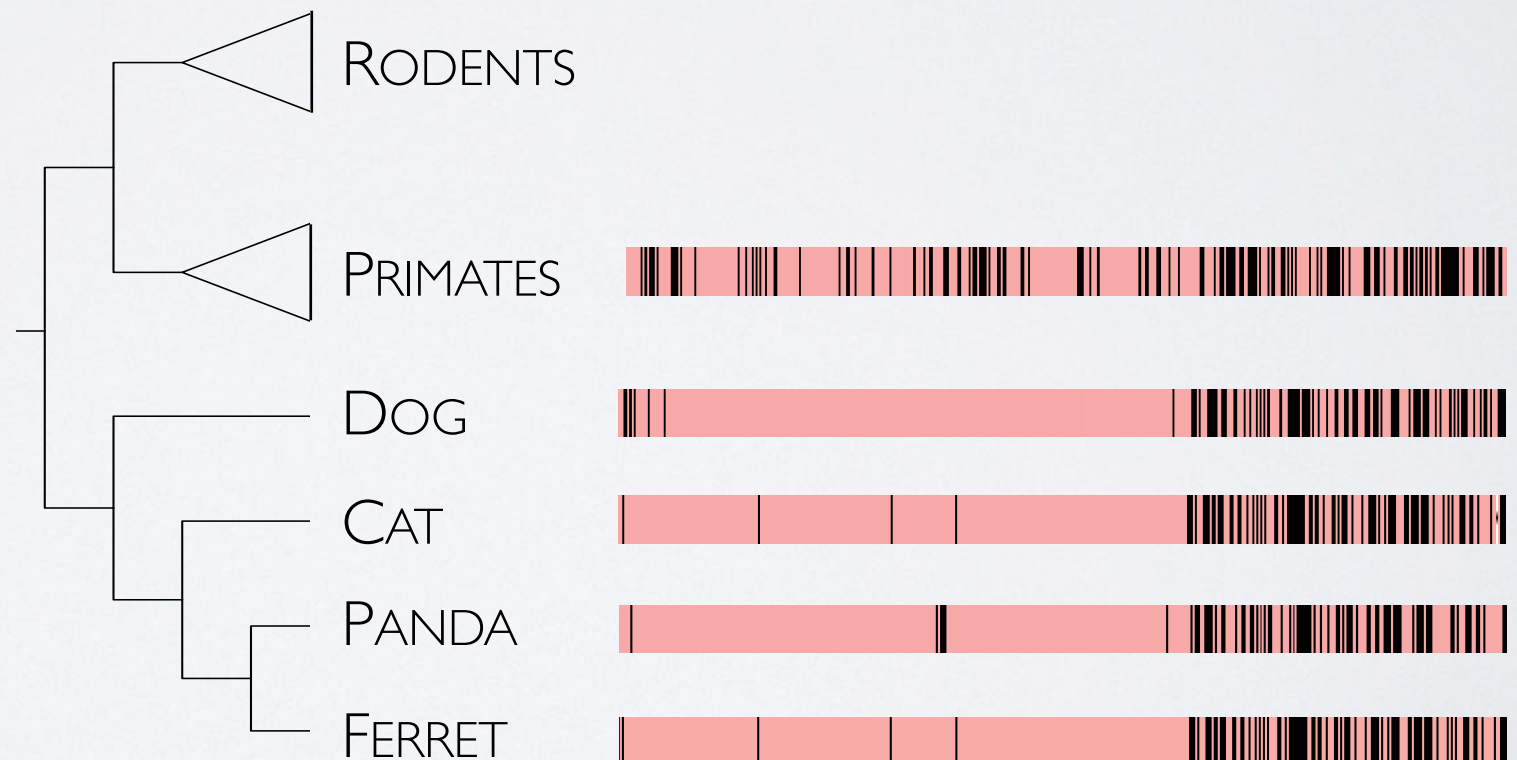
Recurrent recombination among *IFIT1* parallels in mammals



Pairwise comparison of ferret *IFIT1* & *IFIT1B* (black indicates divergence)

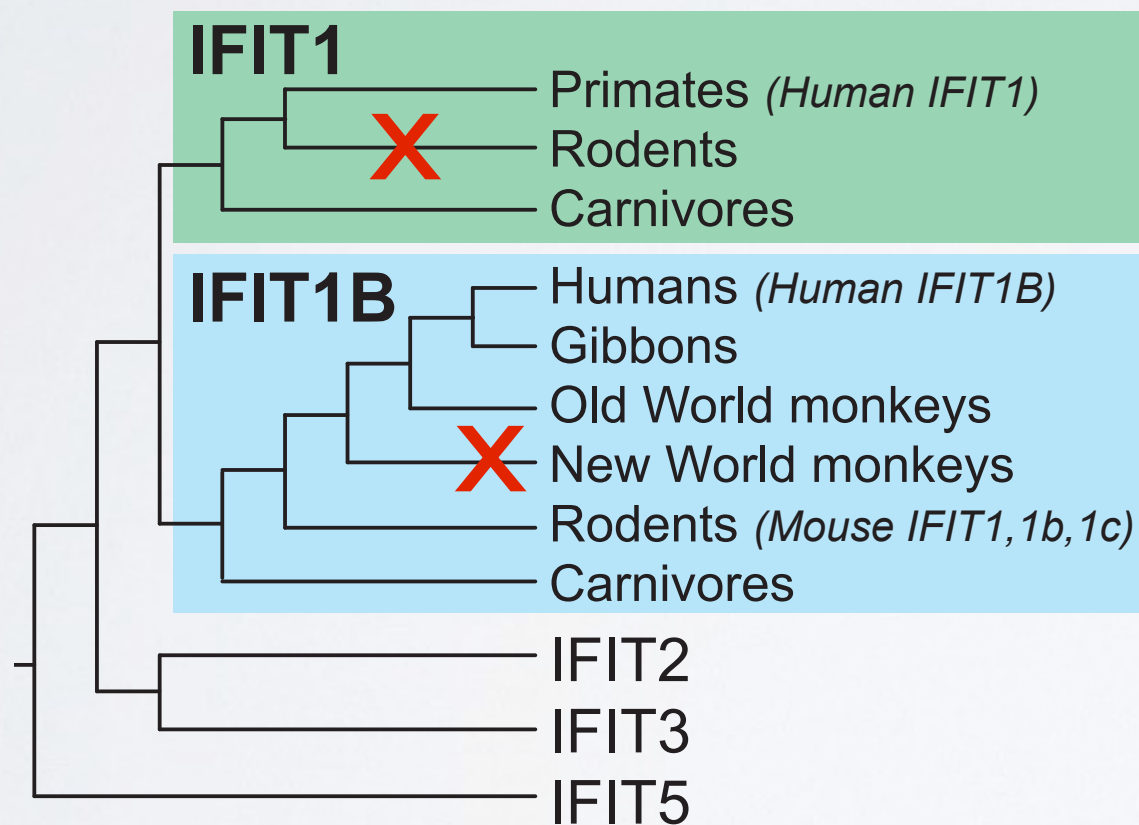
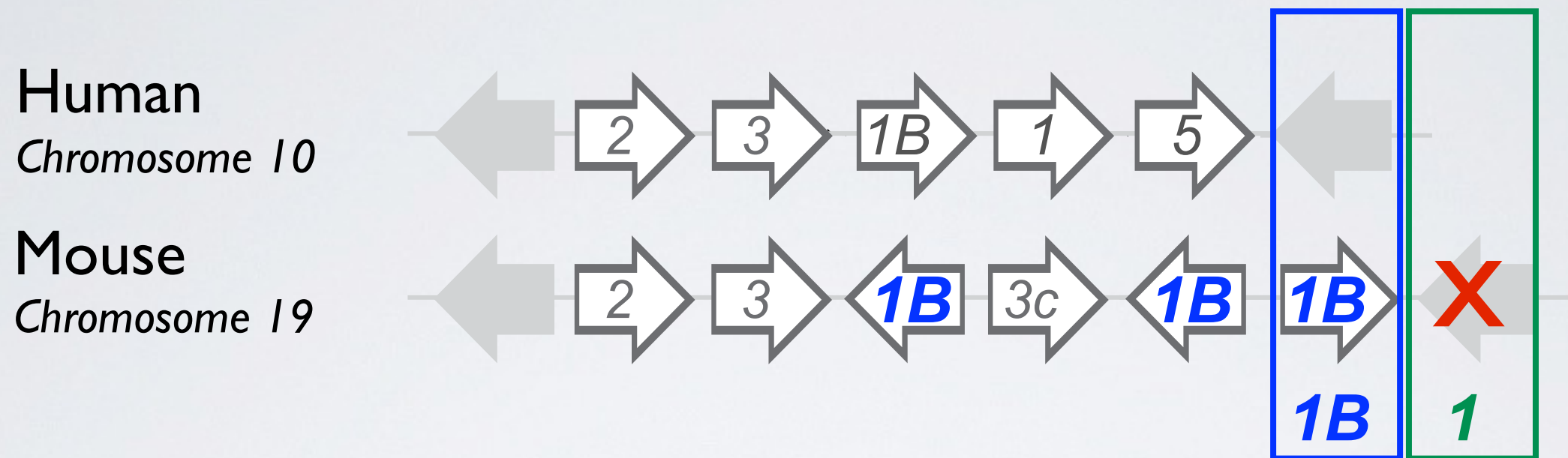


Pairwise comparison of ferret *IFIT1* & Panda *IFIT1*



↑
Identical breakpoint in all 4 carnivores!

TWO ancient families of IFIT1 genes



Missing IFIT1:

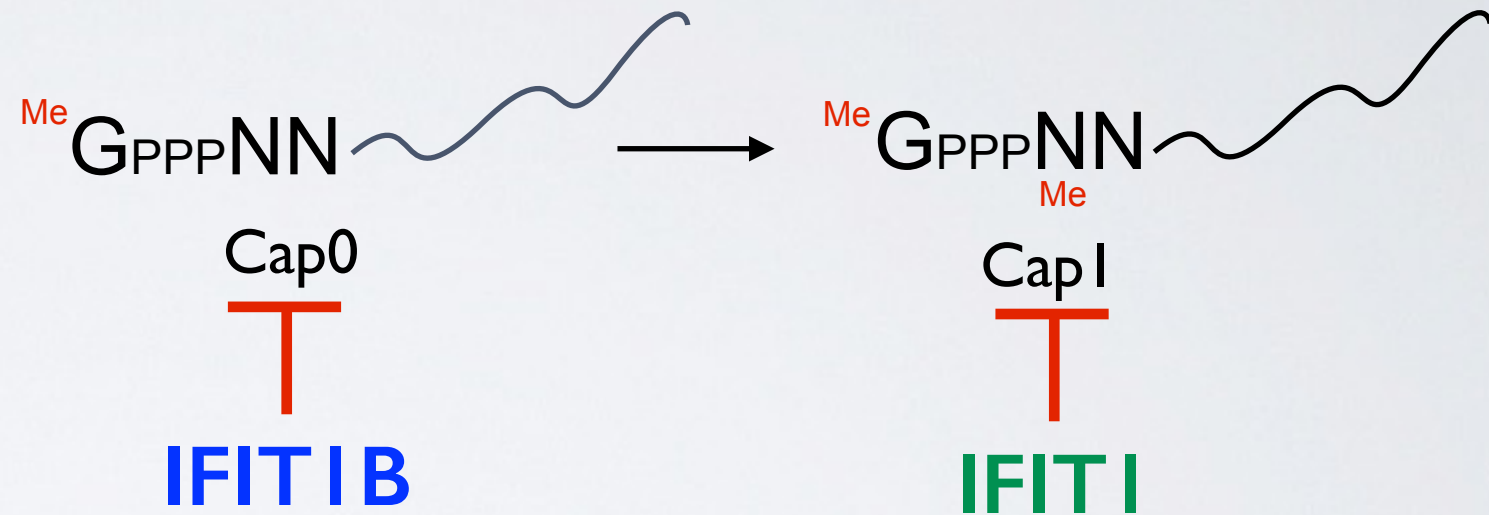
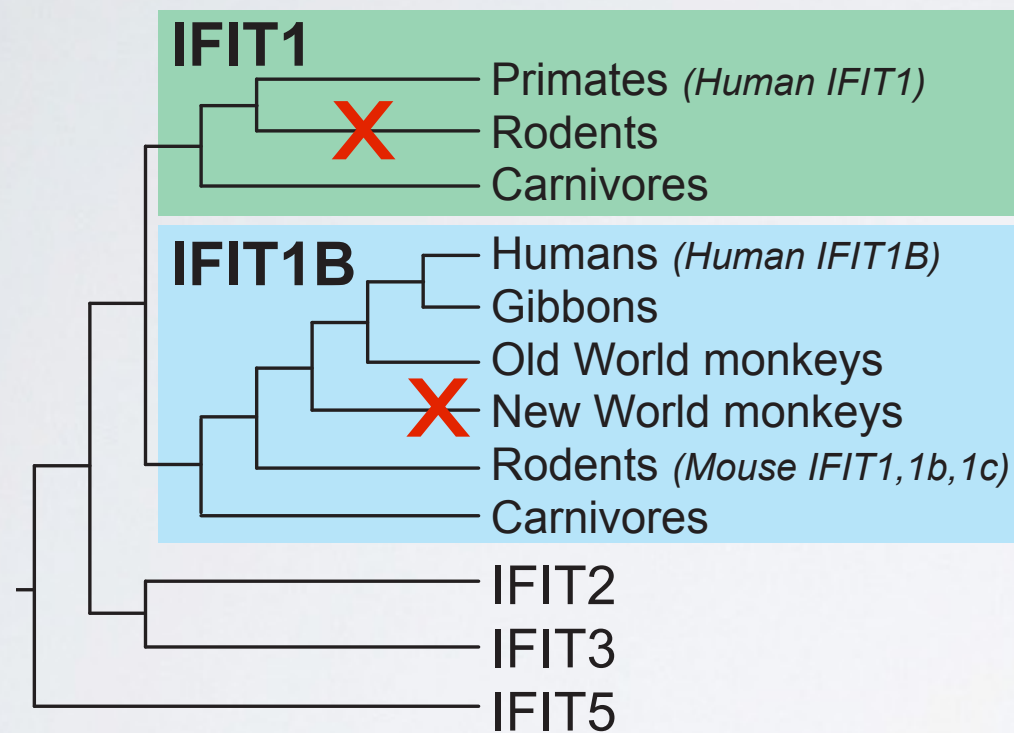
Mice, rats, hamsters, guinea pigs

Missing IFIT1B:

Chimpanzees, New World monkeys,
pigs, cows

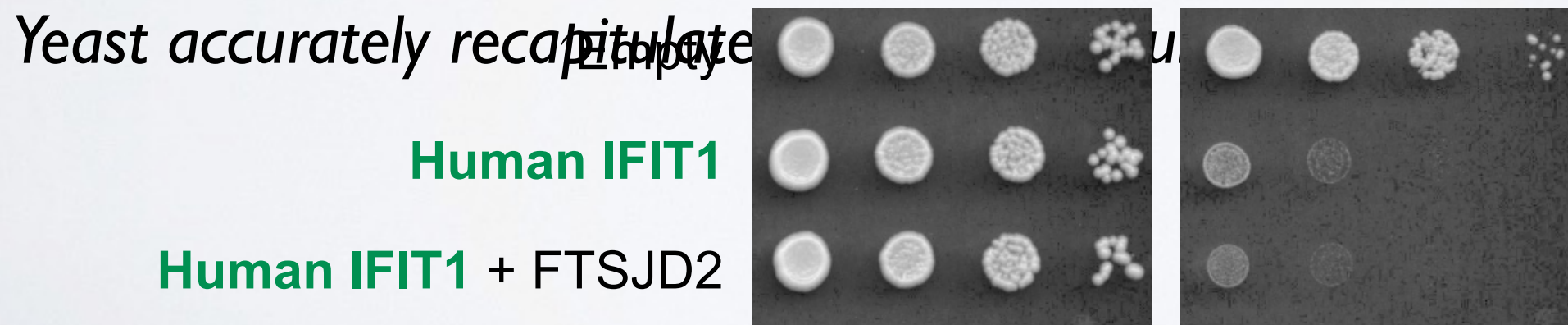
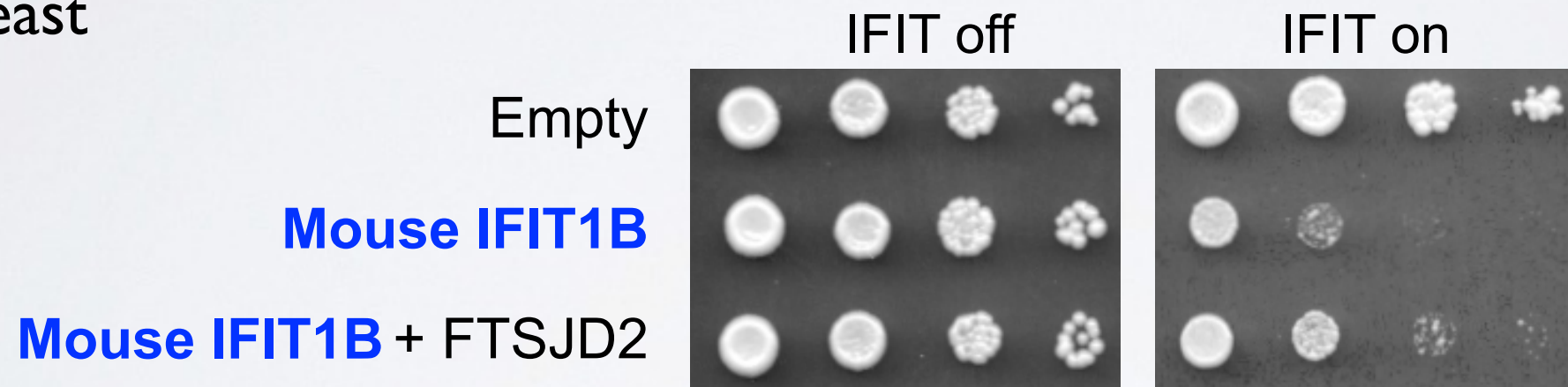
IFIT1 and IFIT1B have been diverging for >100 million years!
But many species only have one or the other!

FUNCTIONAL DIVERGENCE OF IFIT1 PROTEINS?



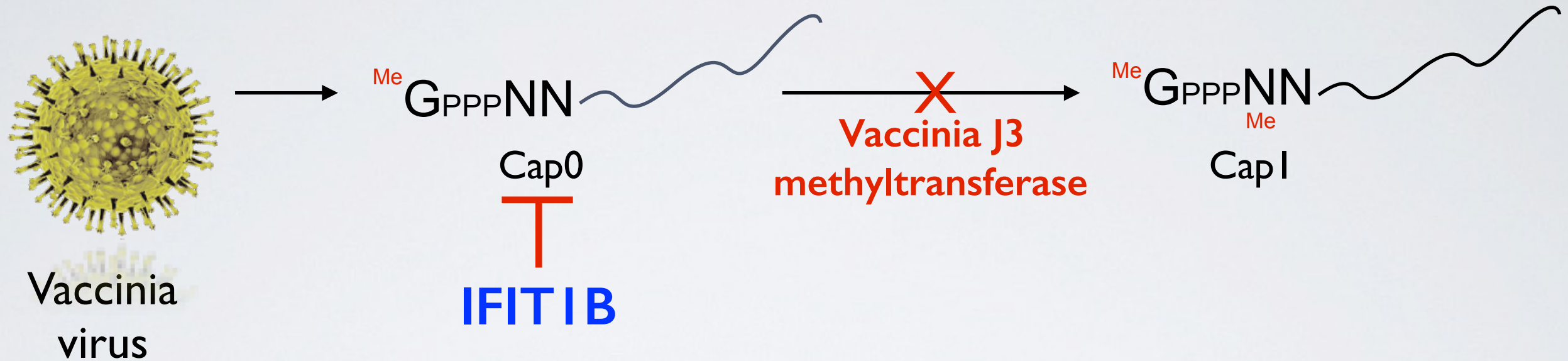
Want to test IFIT activity independent of rest of innate immune system

A NOVEL YEAST ASSAY FOR IFIT FUNCTION

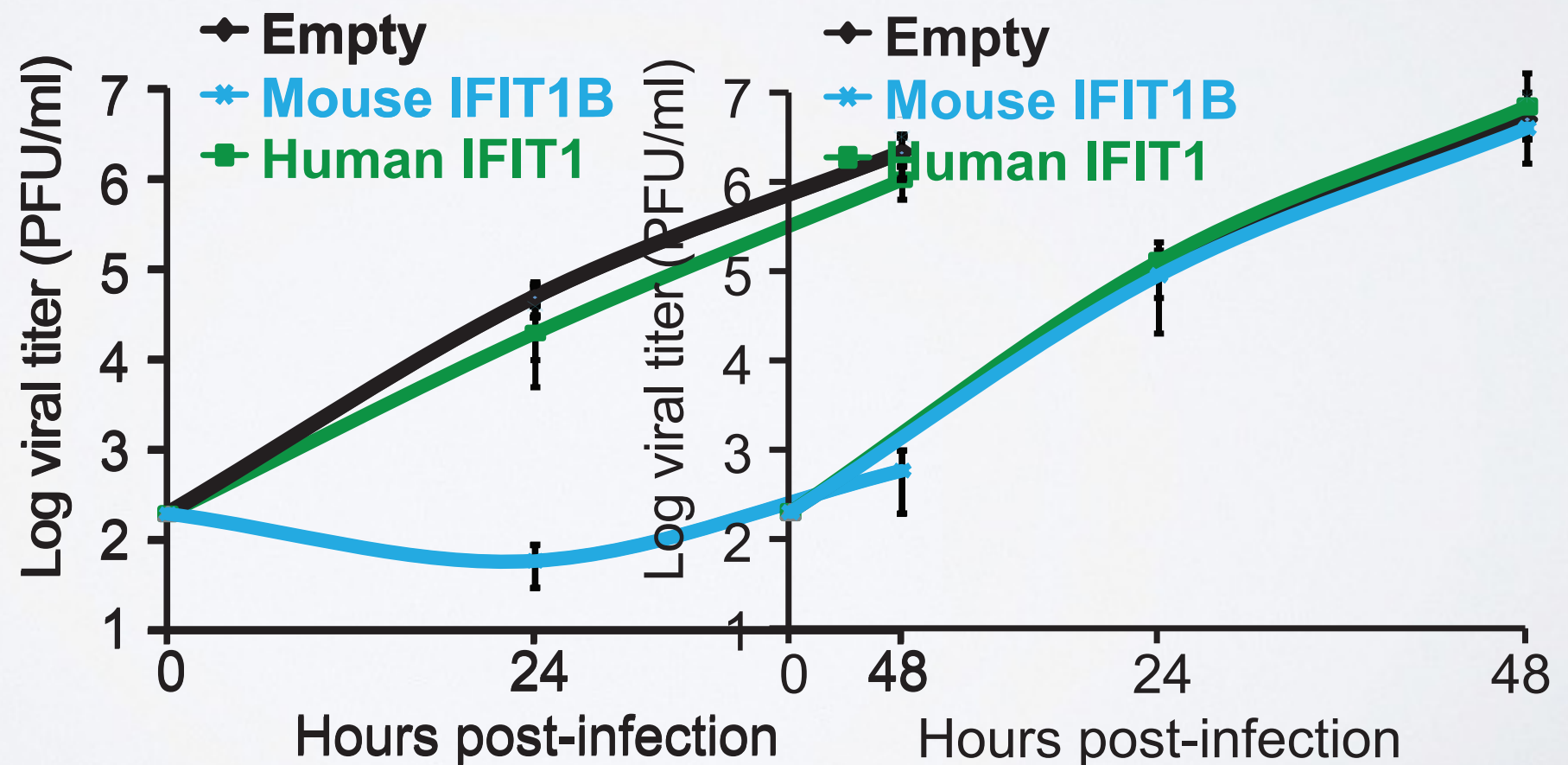


IFIT1 recognizes a different molecular pattern to arrest yeast growth

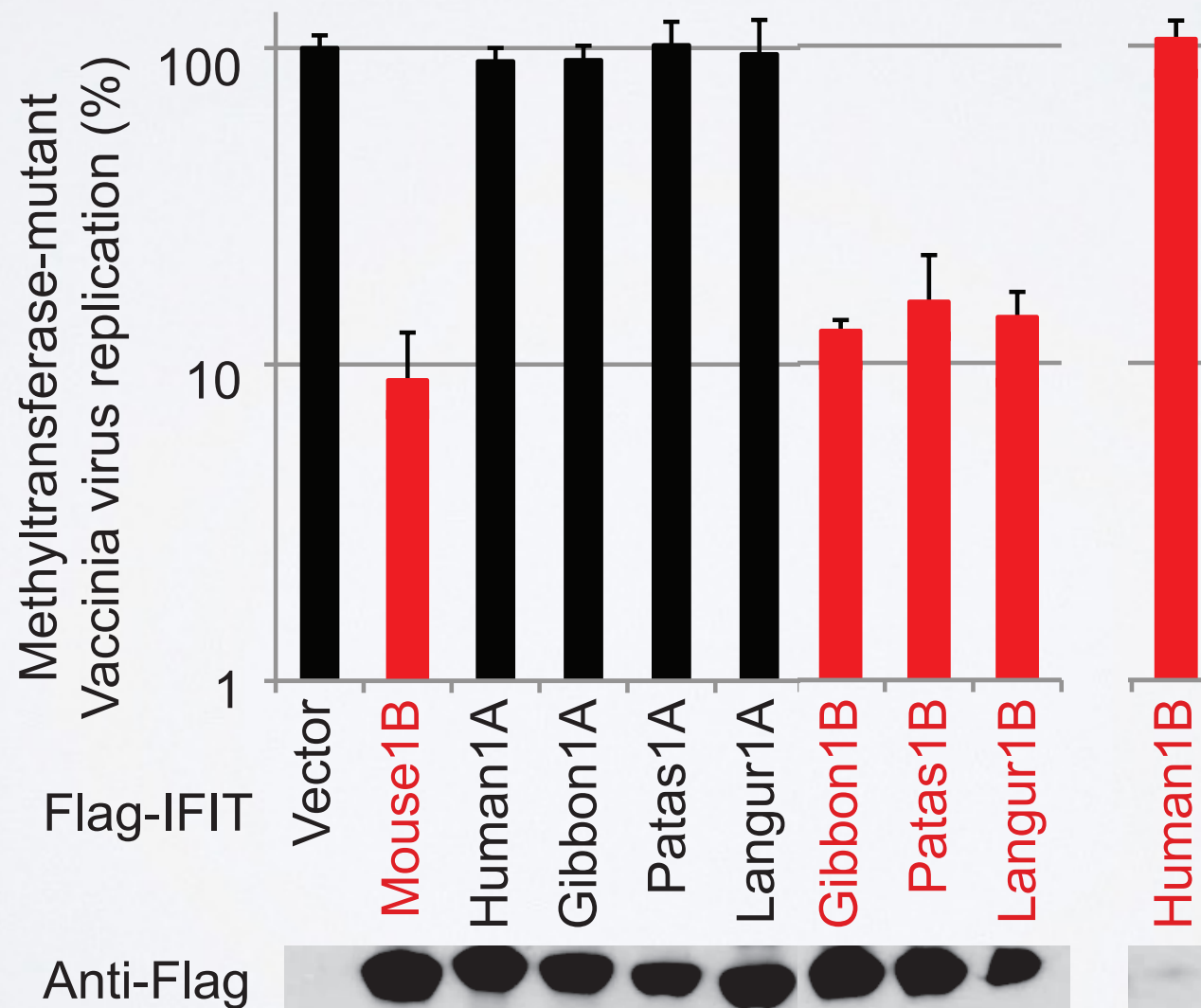
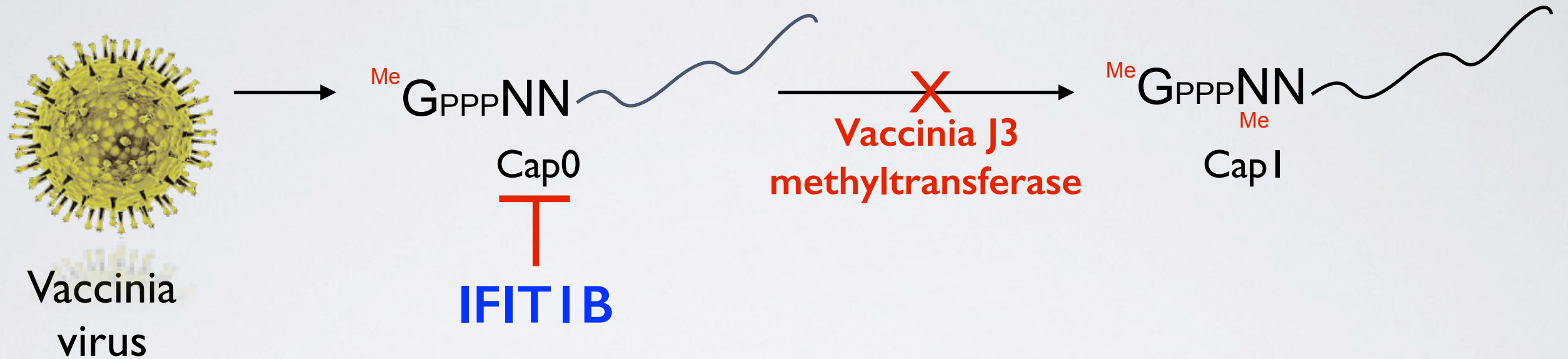
IFIT-MEDIATED ANTIVIRAL RESTRICTION



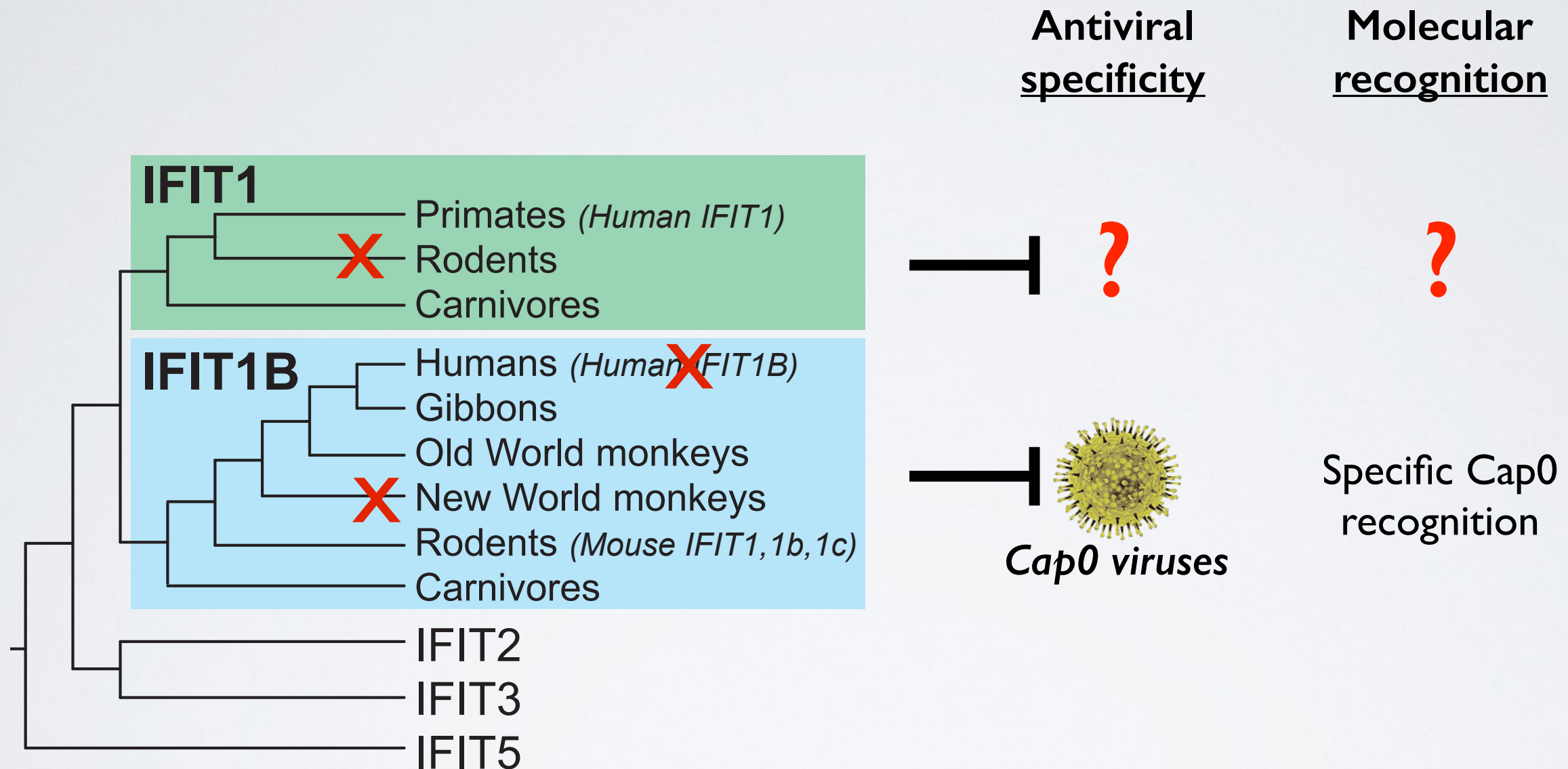
Mutant vaccinia virus (Cap0) vs. vaccinia virus (Cap1)



IFIT-MEDIATED ANTIVIRAL RESTRICTION

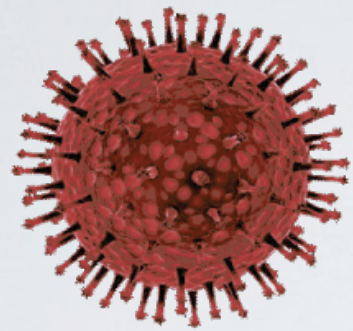


CONSEQUENCES OF IFIT1 EVOLUTION

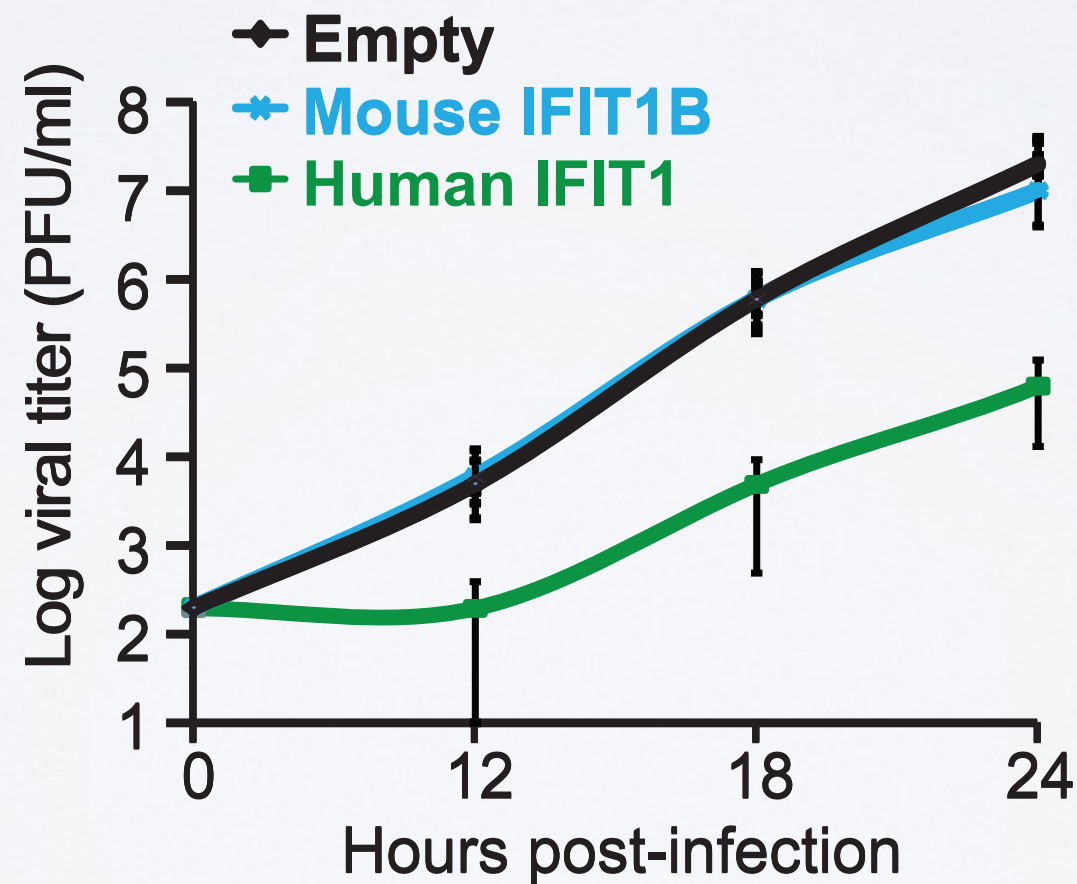
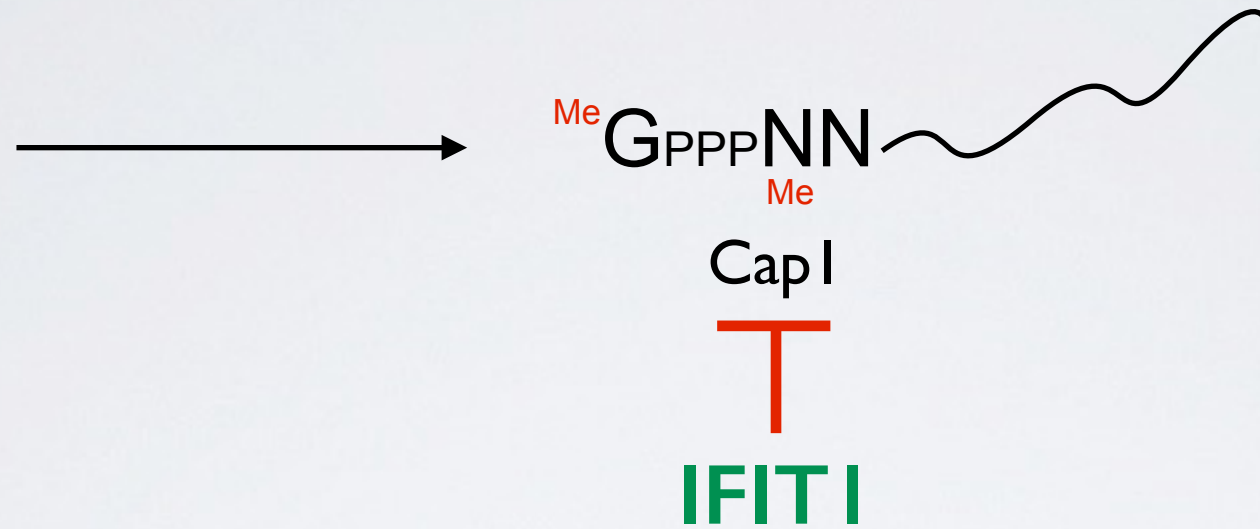


What is the function of IFIT1? Does it have antiviral activity?

IFIT1 INHIBITS CAP1 VIRUS REPLICATION

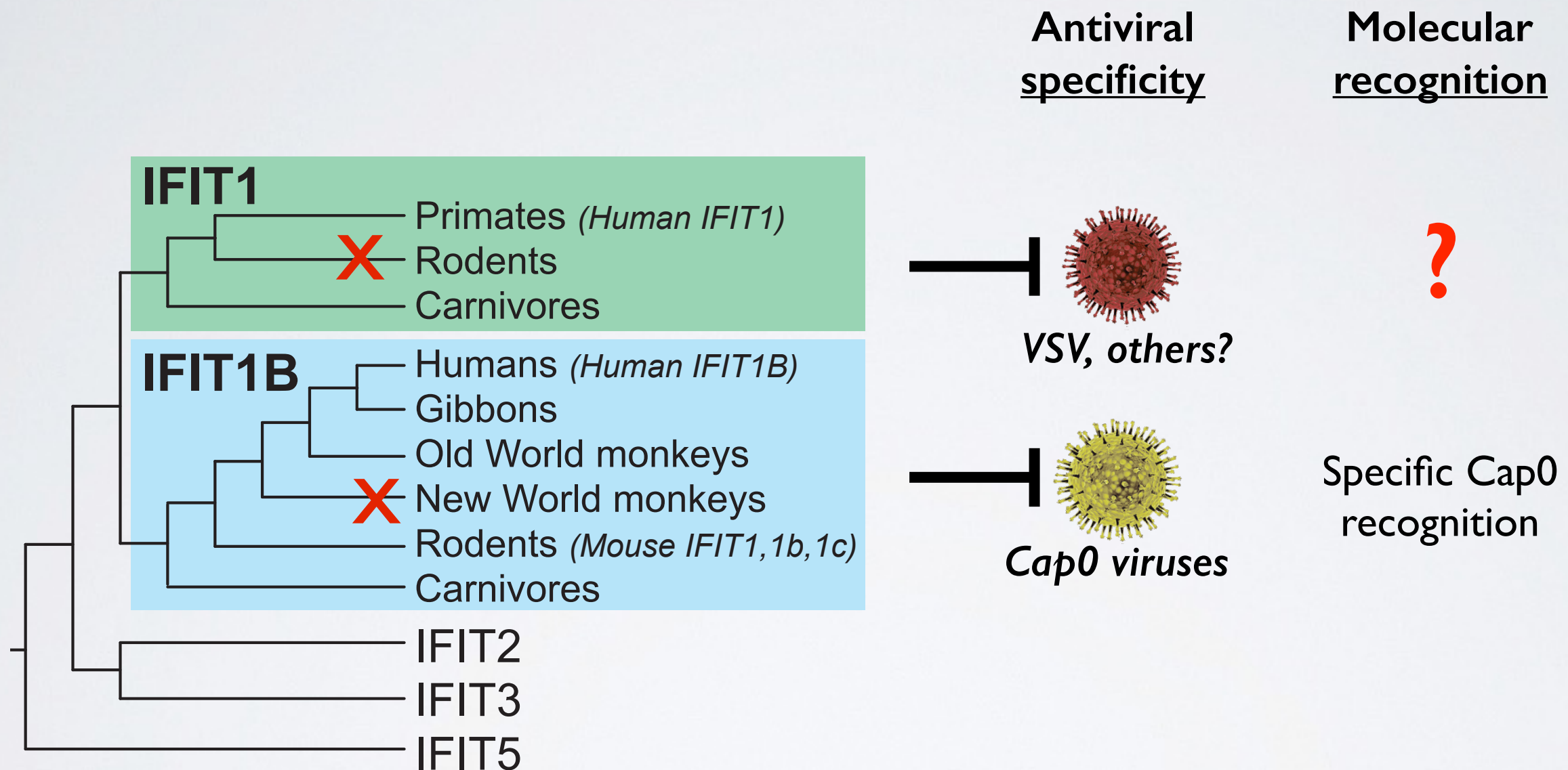


Vesicular
stomatitis
virus (VSV)



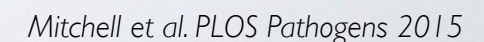
IFIT1 is a potent antiviral, but with specificity that differs from IFIT1B

IFIT1 EVOLUTION ALTERS ANTIVIRAL DEFENSES

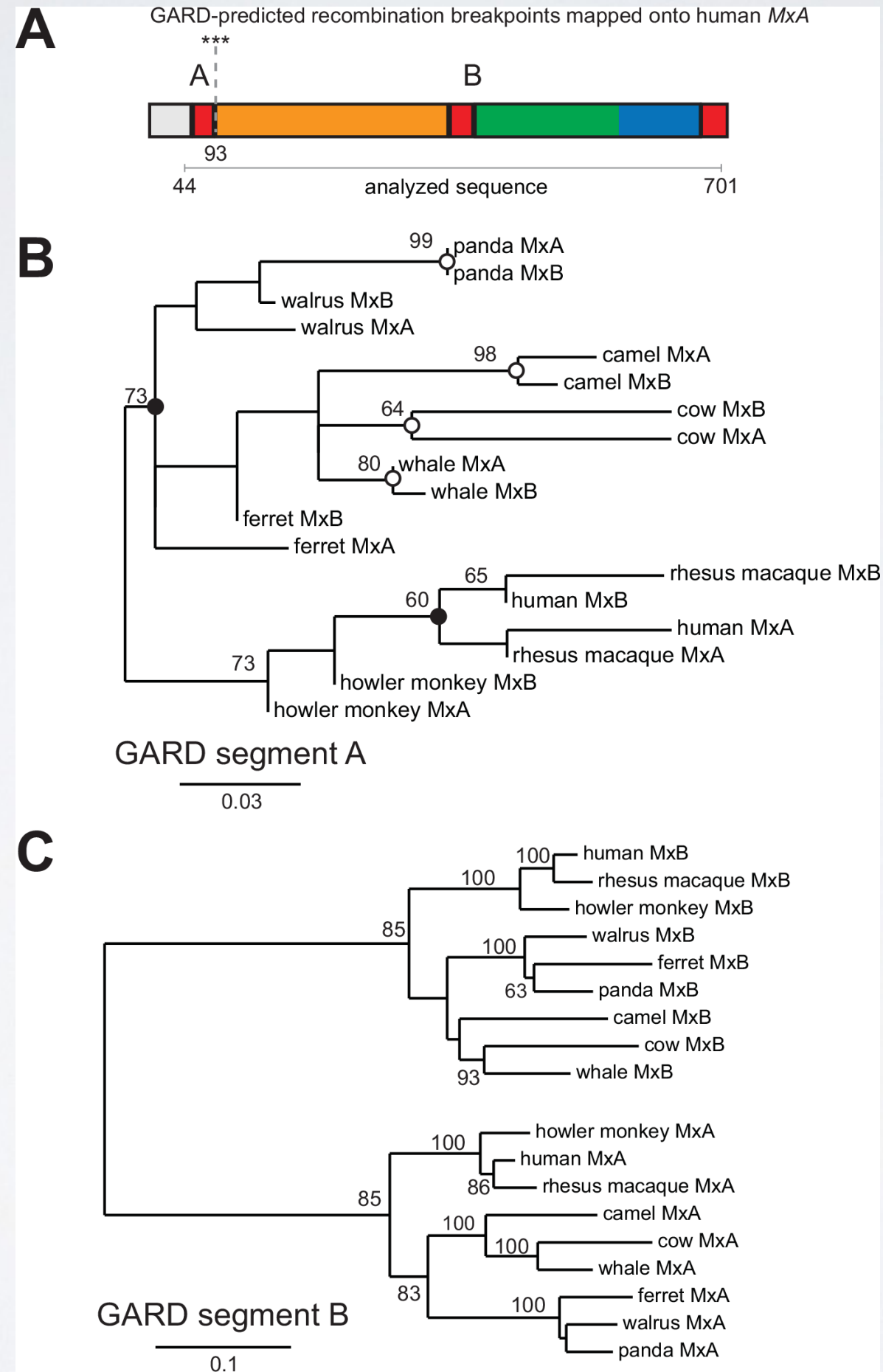


IFIT1 and IFIT1B are distinct gene families that diverged millions of years ago
IFIT1 and IFIT1B have evolved different mechanisms and antiviral specificities

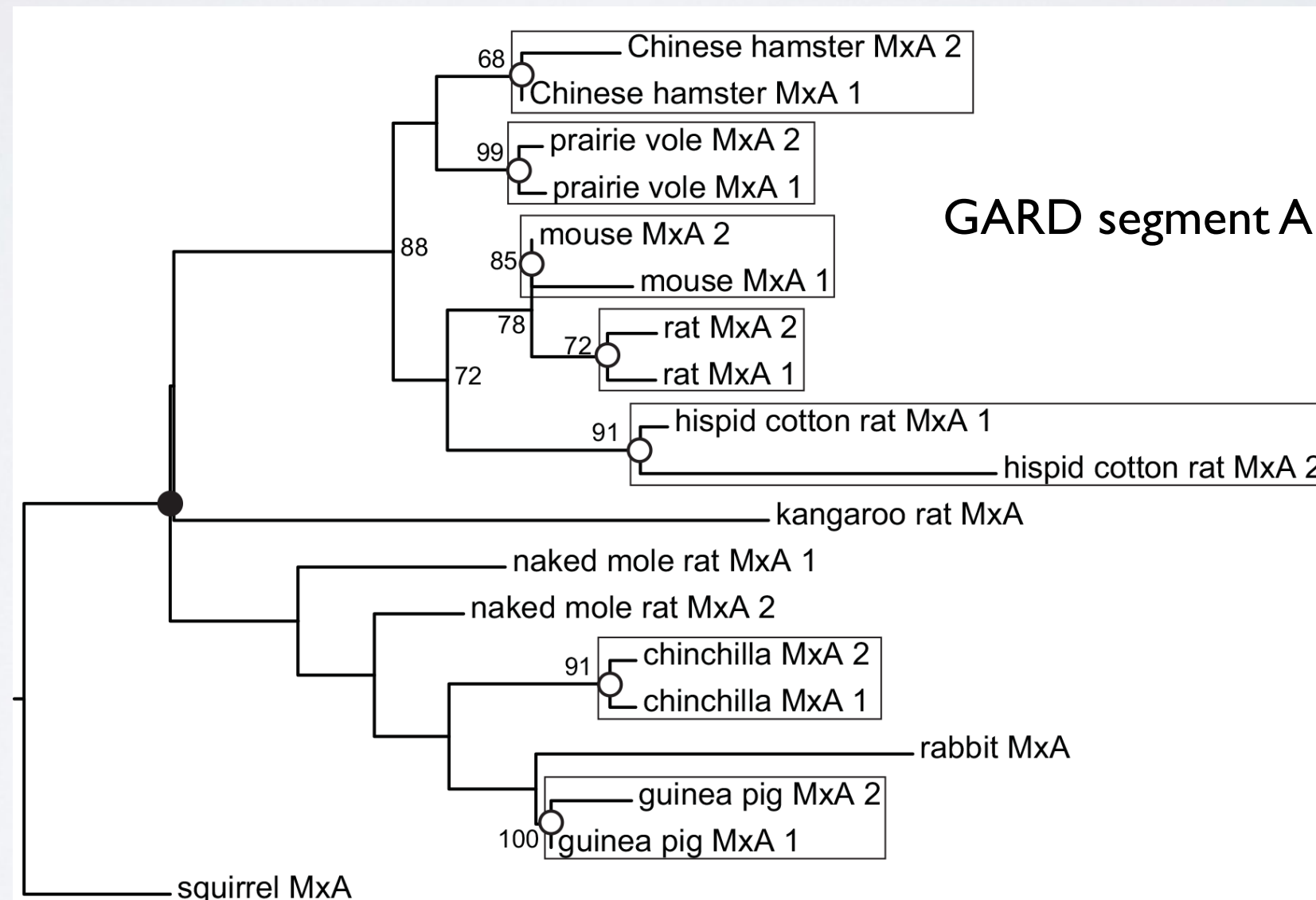
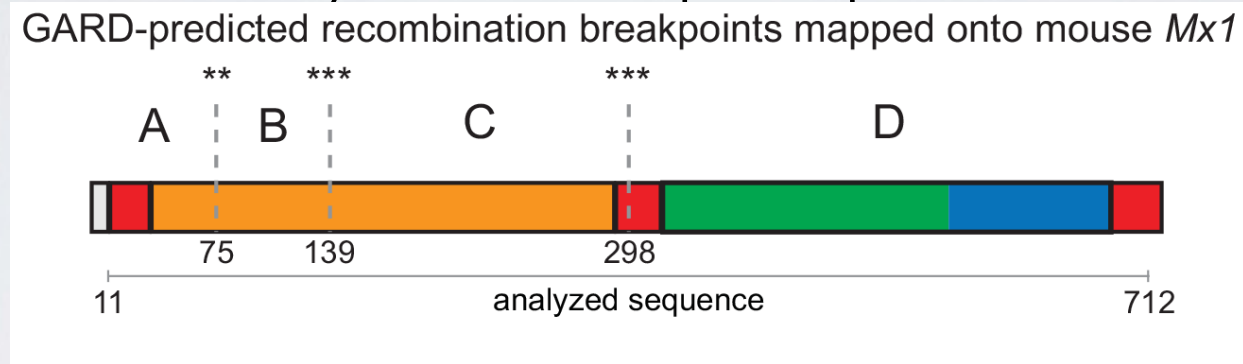
Dr. Yoon is a smiling man with dark hair and a goatee, wearing a white lab coat over a dark shirt. He is standing in a laboratory setting with shelves of equipment and supplies in the background. A small logo is visible in the bottom right corner of the photo.



Recombination may also help explain MxA/MxB evolution

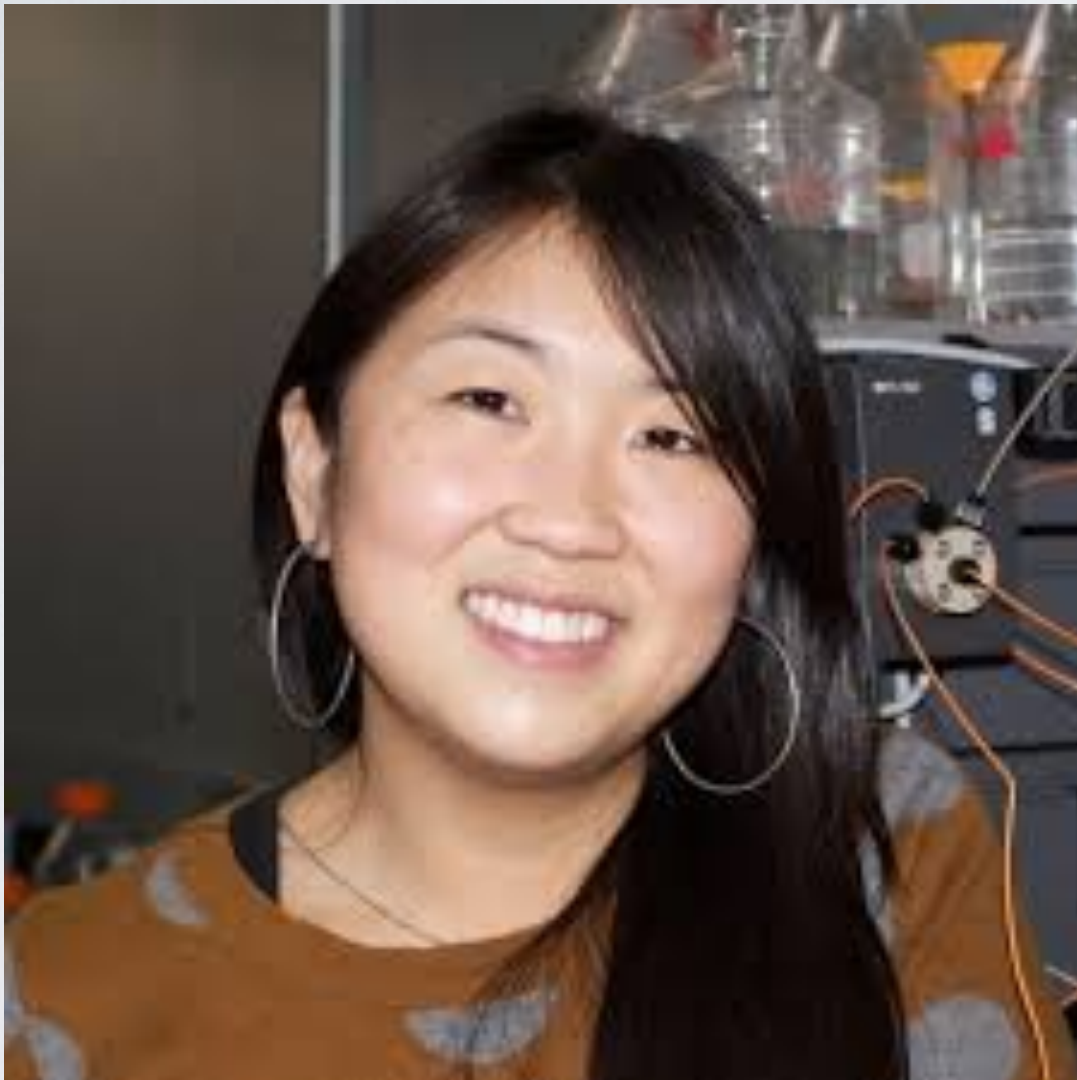


Recombination may also help explain MxA/MxB evolution



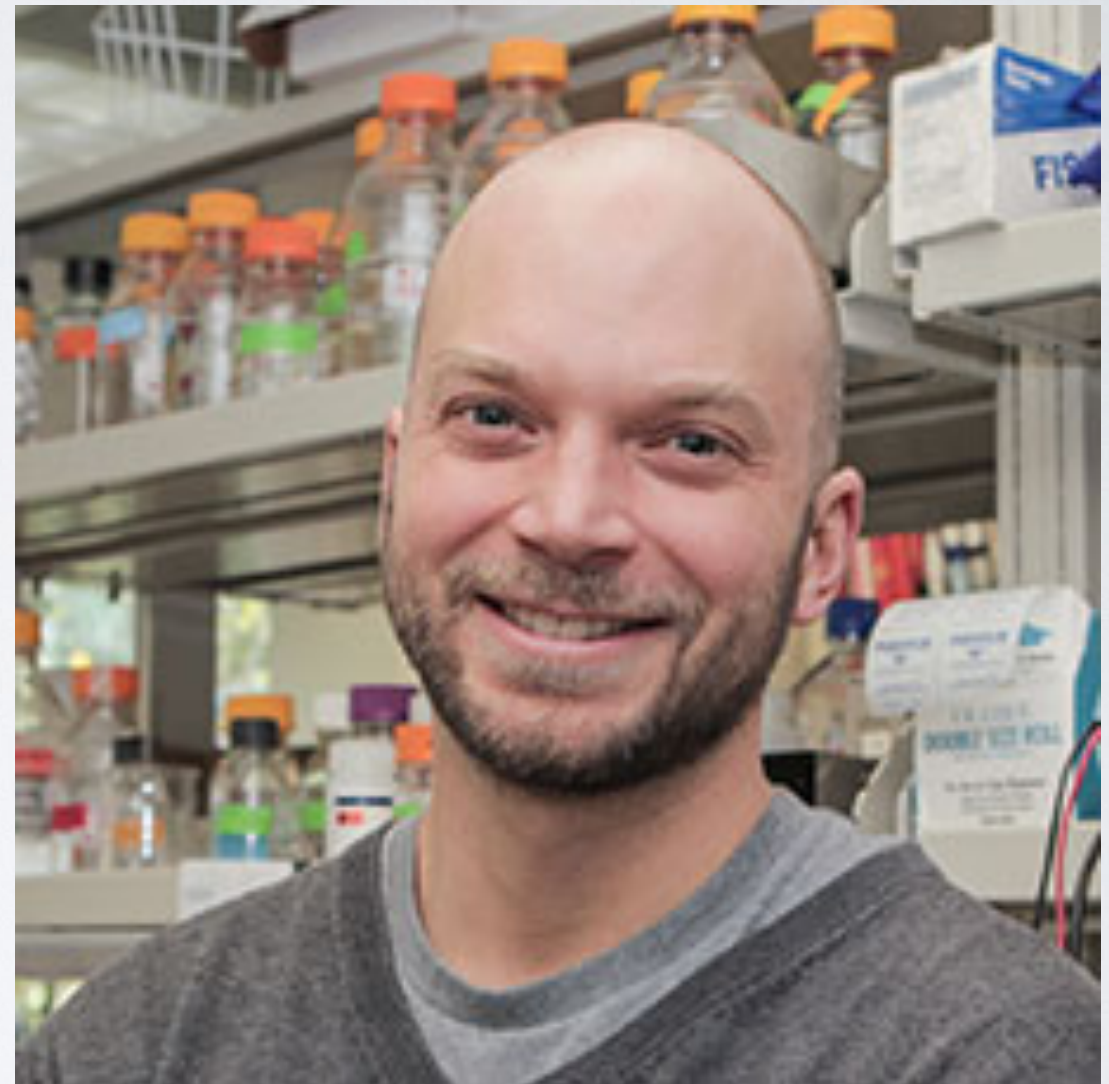
Lecture 3- Vignette 3

An unusual origin for a host immune gene



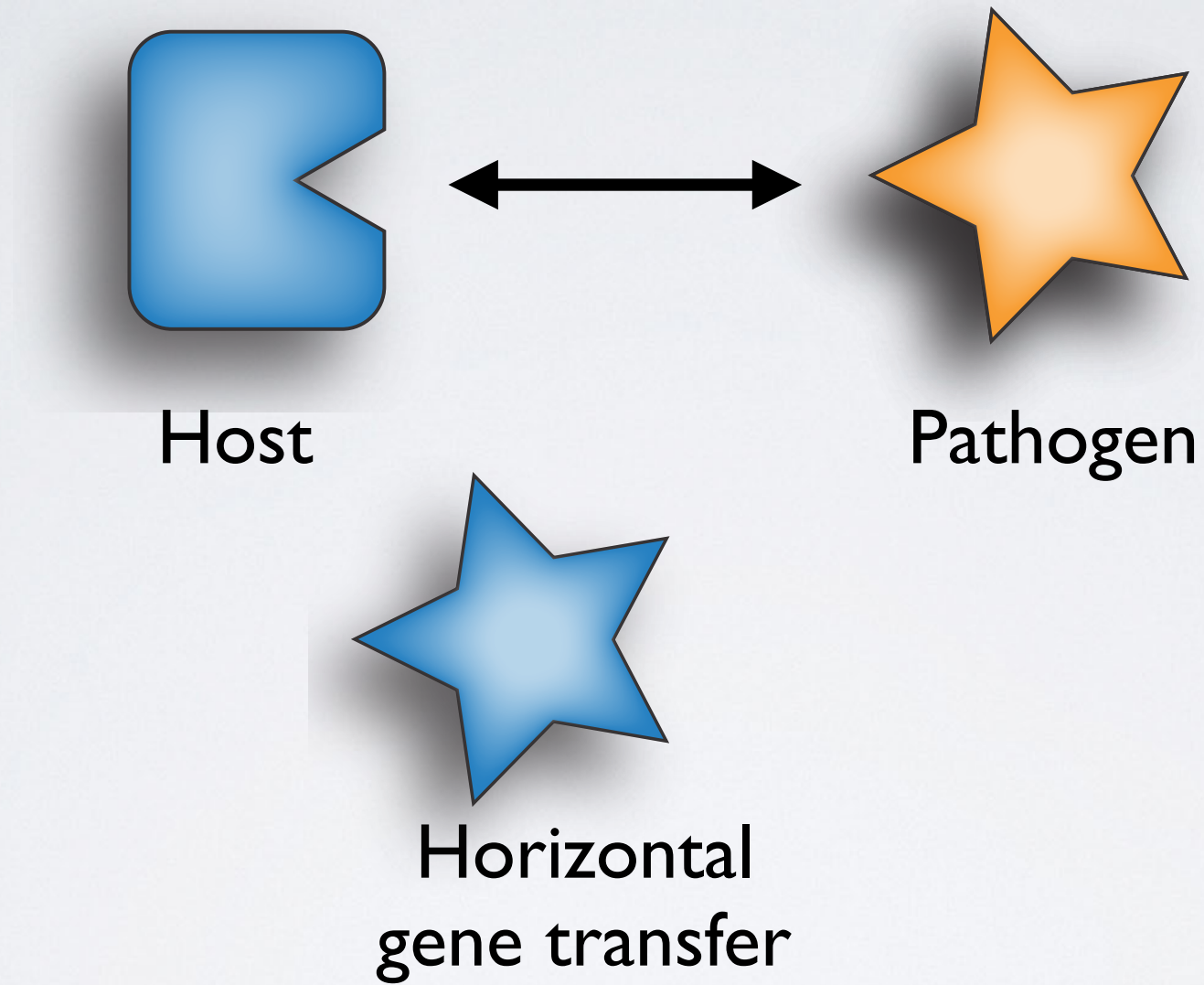
Seemay Chou

former postdoc, Mougous lab, UW
now Assistant Professor, UC San Francisco

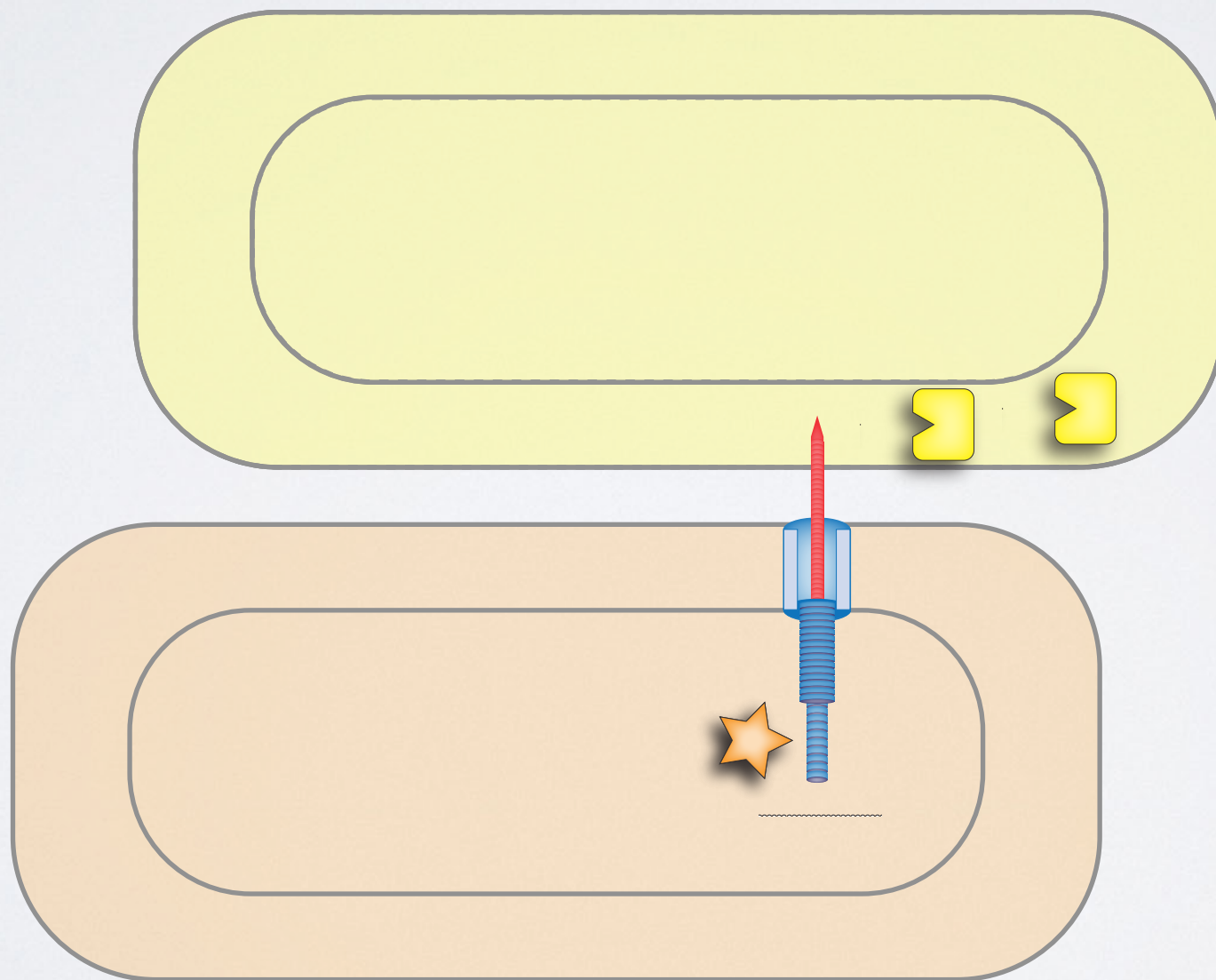
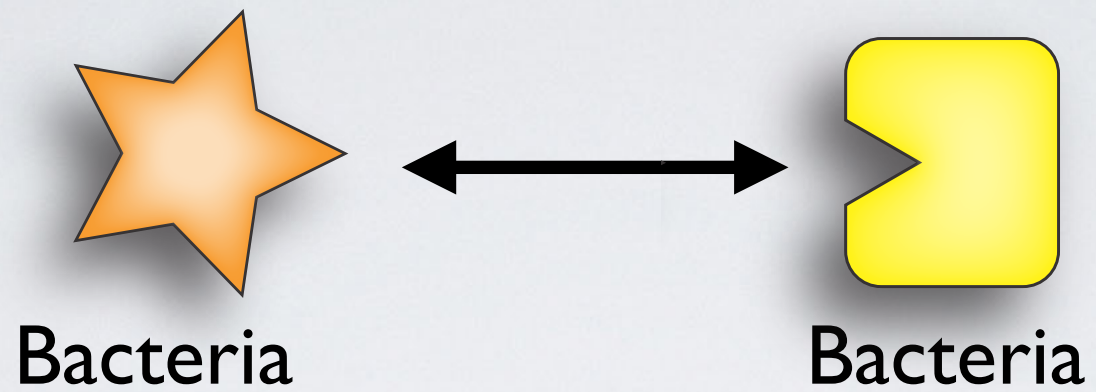


Matt Daugherty, former postdoc
now Assistant Professor, UC San Diego

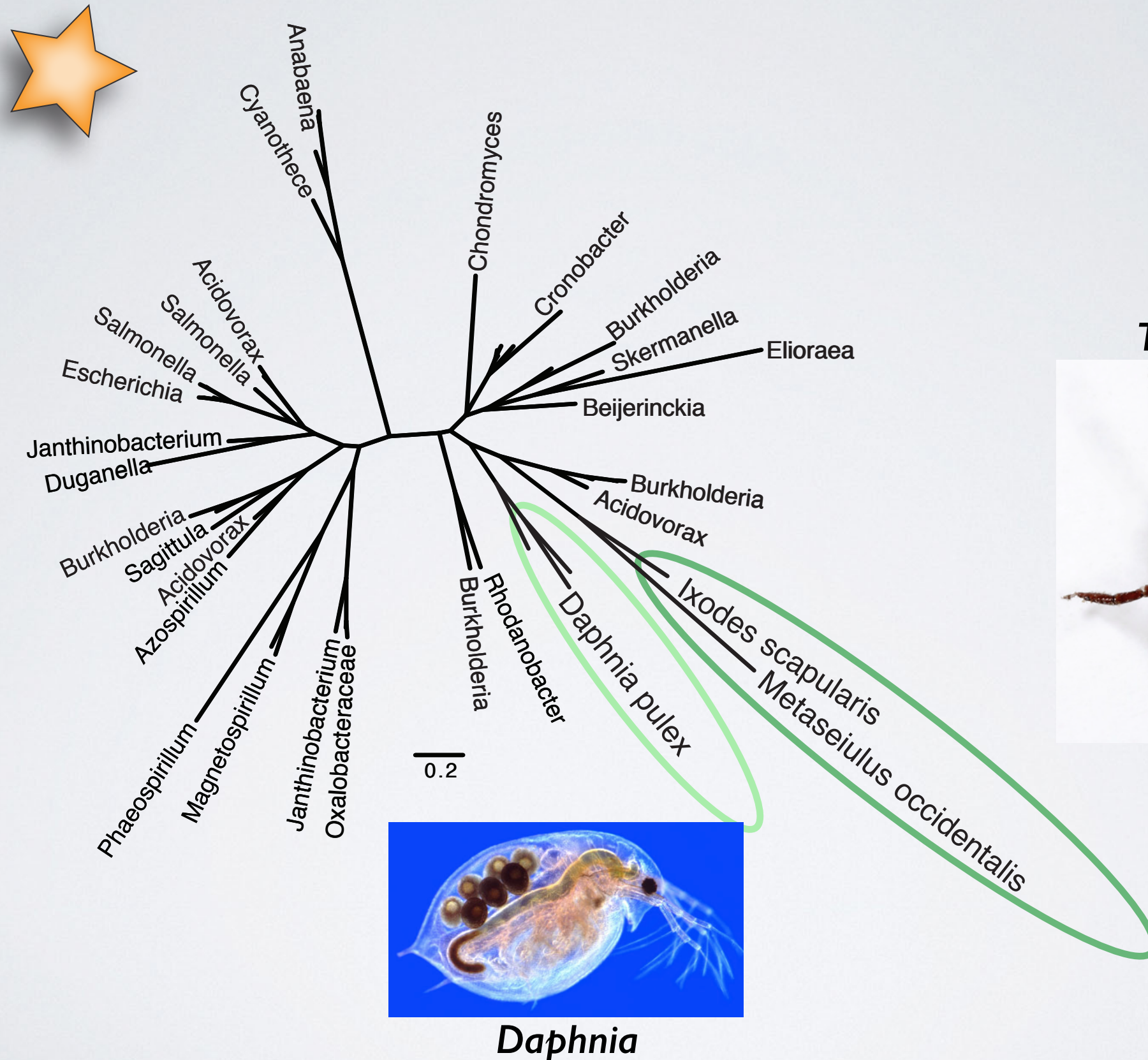
'Stealing' immunity genes



Interbacterial warfare



Bacterial genes in eukaryotic genomes



Ticks and mites



Daphnia

Antibacterial effector genes were 'stolen' at least 6 times

Ticks and mites



Oxytricha



Mollusks



Daphnia



Naegleria

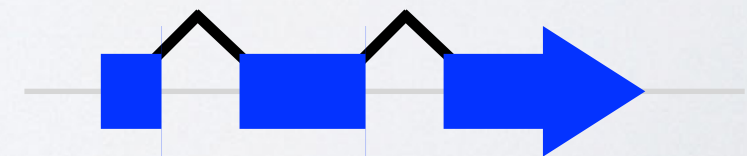


Sea anemone, Lancelet and others

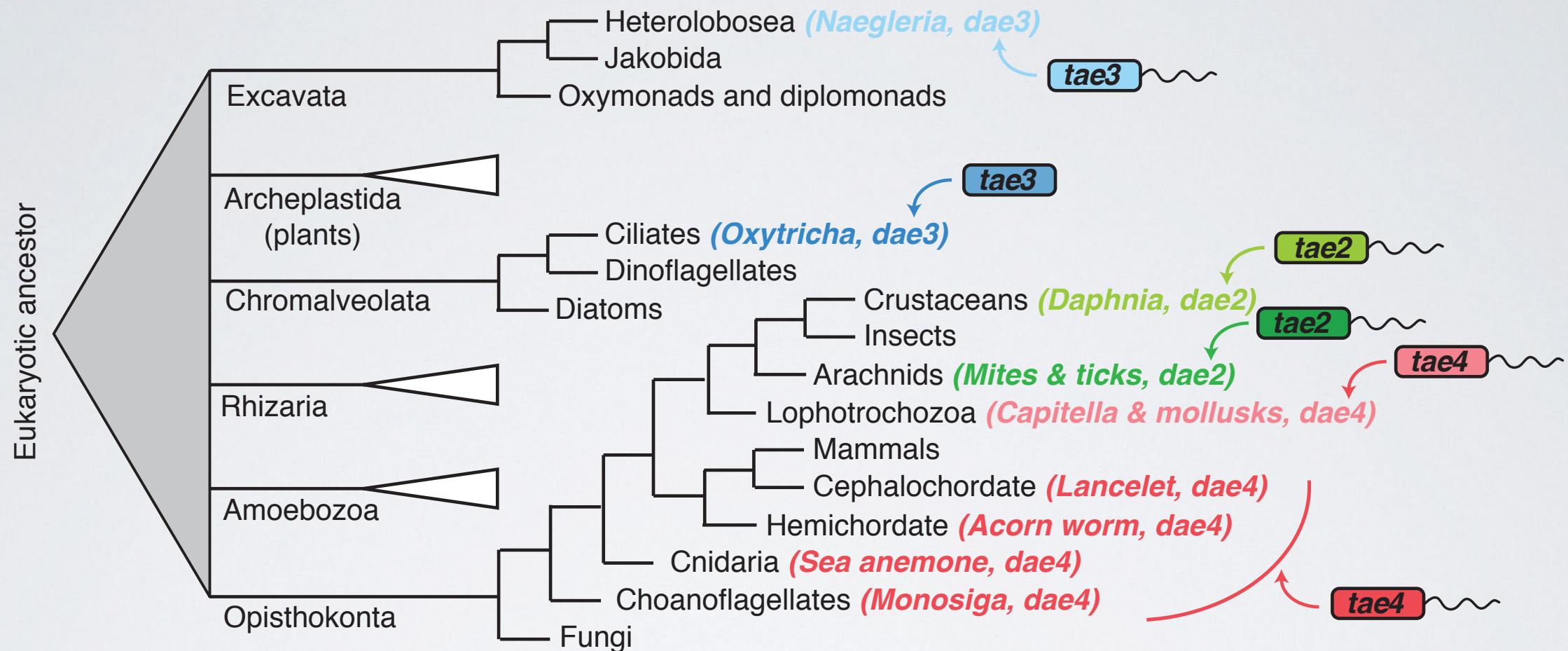
Surrounded by other eukaryotic genes



Contain introns



GENES WERE 'STOLEN' AT LEAST 6 TIMES



Renamed **Domesticated amidase effector (*Dae*)** genes

Conservation of catalytic residues

... and a new way to get out of cells

		Conserved catalytic residues		New secretion signal
<i>Bacterial consensus</i>				
>400 million years of tick and mite evolution	<i>A. cajennense</i>	KE C VALVK	H AAIFES	MKGFVICAALLVLGMAVGSSQAI
	<i>A. triste</i>	KE C VALVK	H AAIFES	MNGFLLCAALLVLGMAVGSSQAI
	<i>A. parvum</i>	KE C VALVK	H AAIFES	MKCFVTCTALLVLGMAVGSSQAI
	<i>A. maculatum</i>	KE C VALVK	H AAIFES	MNGFILSAALLVLGMAVGSSQAI
	<i>R. pulchellus</i>	RE C VALVK	H AAIFES	MKGFVVSVGLLLLGMIAIVIQA I
	<i>R. sanguineus</i>	KE C VALVK	H AAIFES	MKGFVISVGLLLLGMTVVTQA I
	<i>I. scapularis</i>	KE C VALVK	H AAIFES	MKMGVTPLVLVVSIAFSLNCEVD A I
	<i>M. occidentalis</i>	GSAAALIK	H AGIFVR	MKLFLISAALVVLGLAAVSDAM
	<i>D. gallinae</i>	GSAAGLVK	H AAILVR	MSPSTVTGALLALLAFTVGVSGM
<i>D. pteronyssinus</i>	QE C AALVQ	H TAIFVR	MYSFKYCLAAAILLLVGNVCQA F	

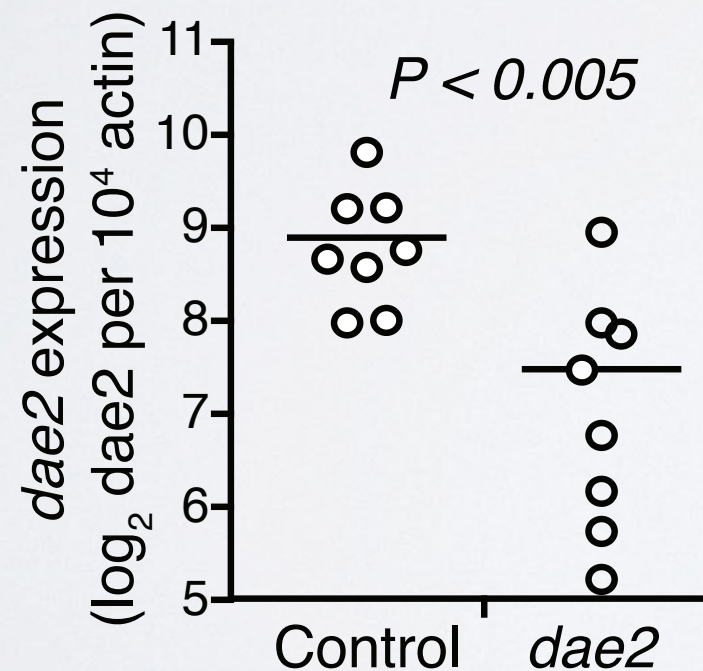
Dae-mediated antibacterial immunity *in vivo*



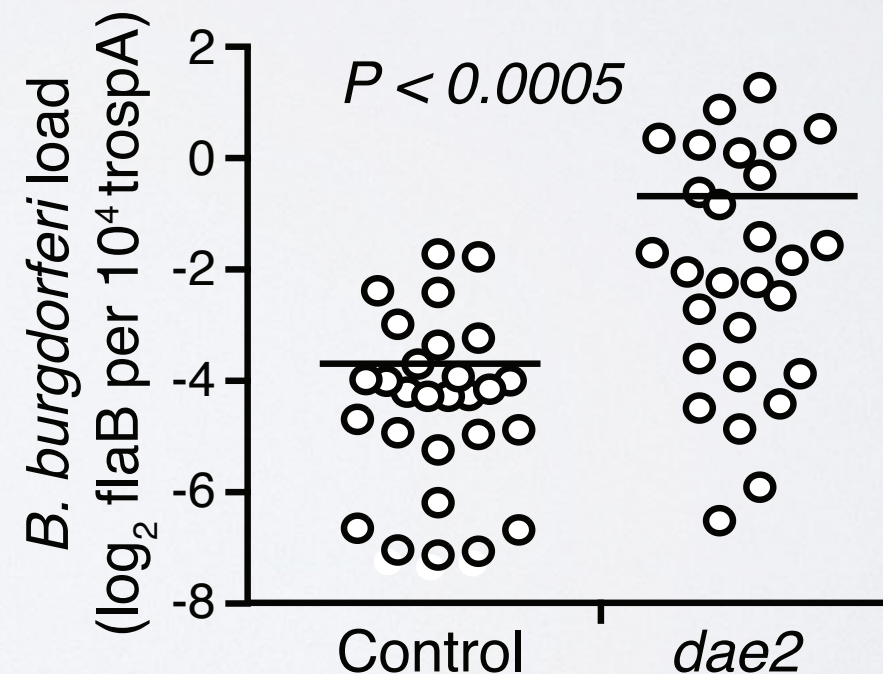
Ixodes scapularis (deer tick) is the vector for the Lyme disease pathogen (*Borrelia burgdorferi*)

Lyme disease is the most prevalent vector-borne illness in the US

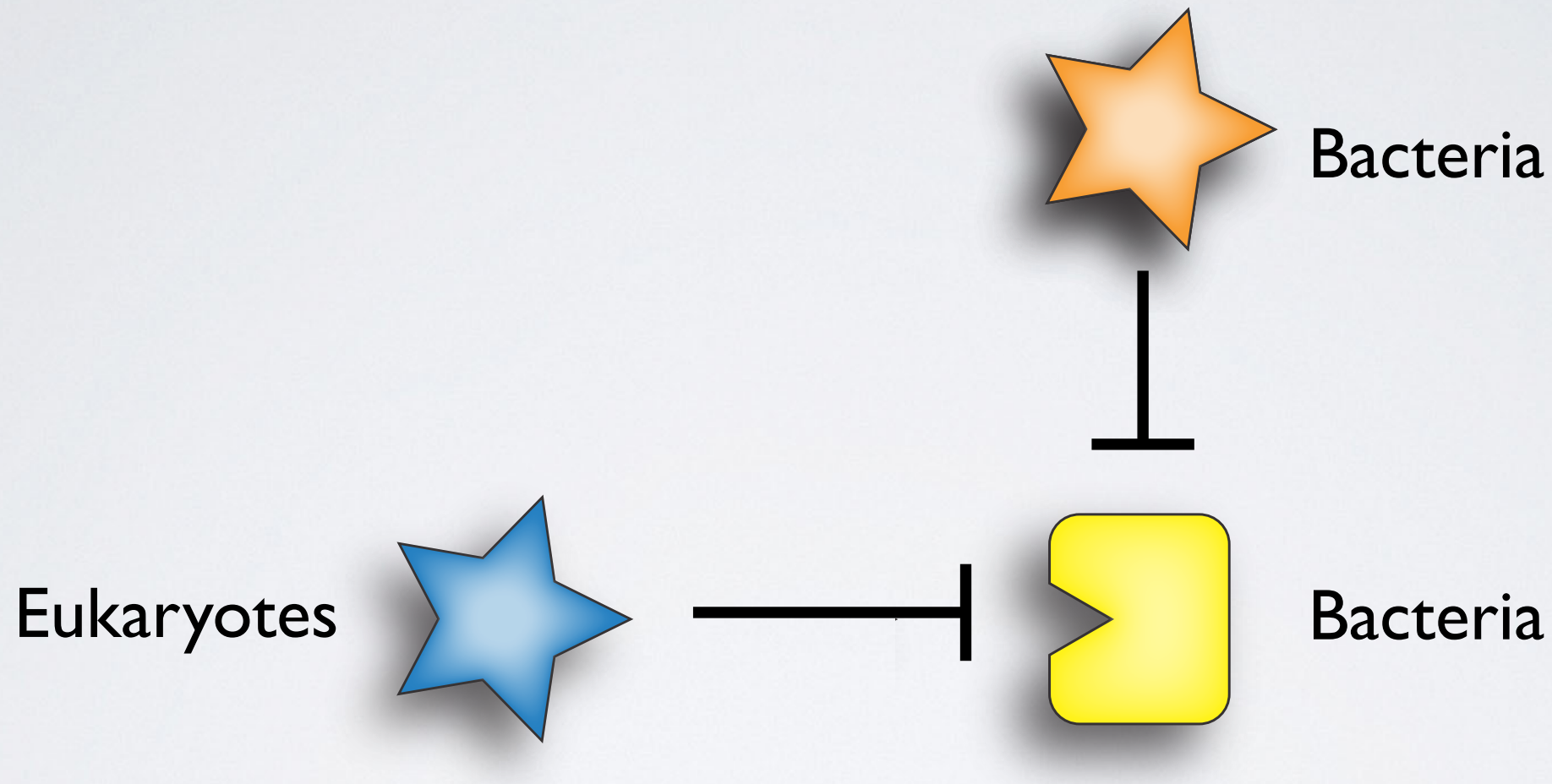
RNAi knockdown



2 weeks post-engorgement

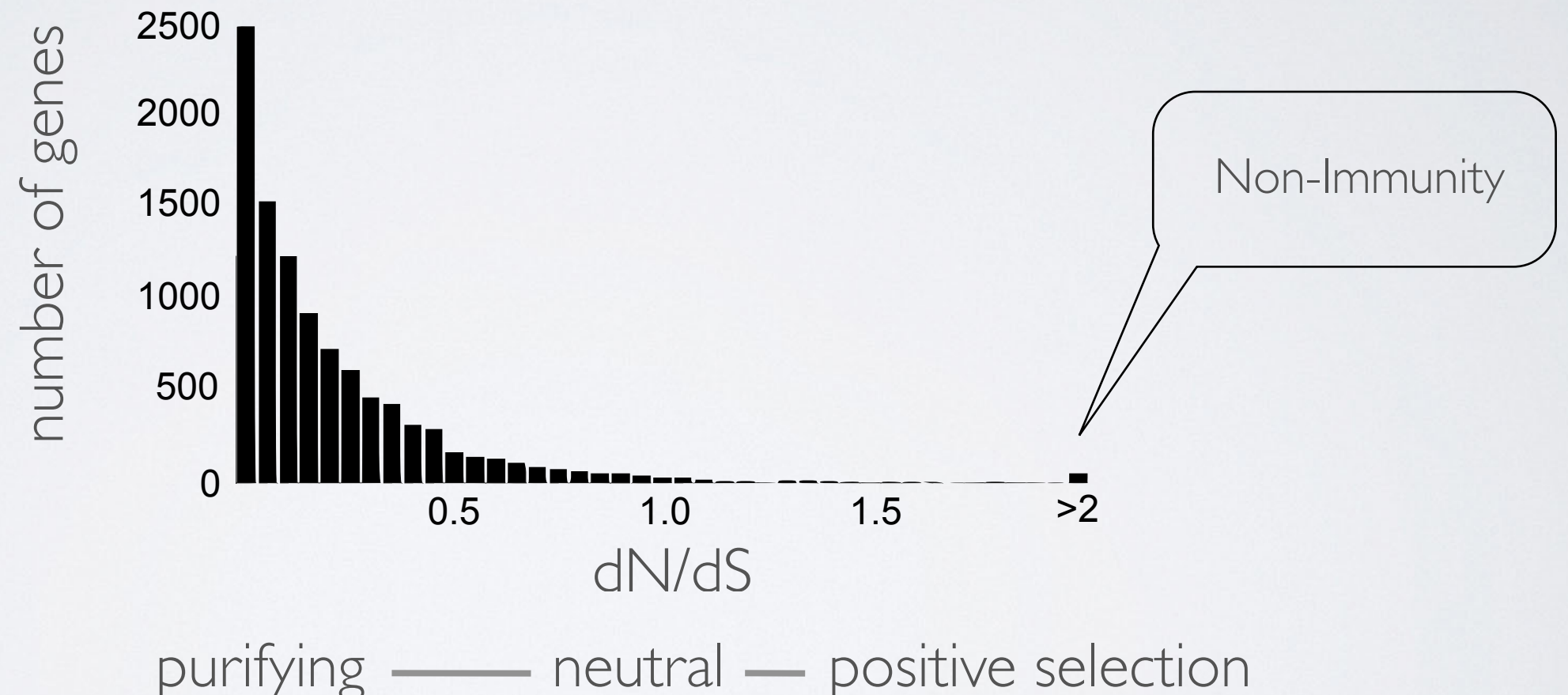


Horizontal gene transfer as a mechanism to acquire immunity



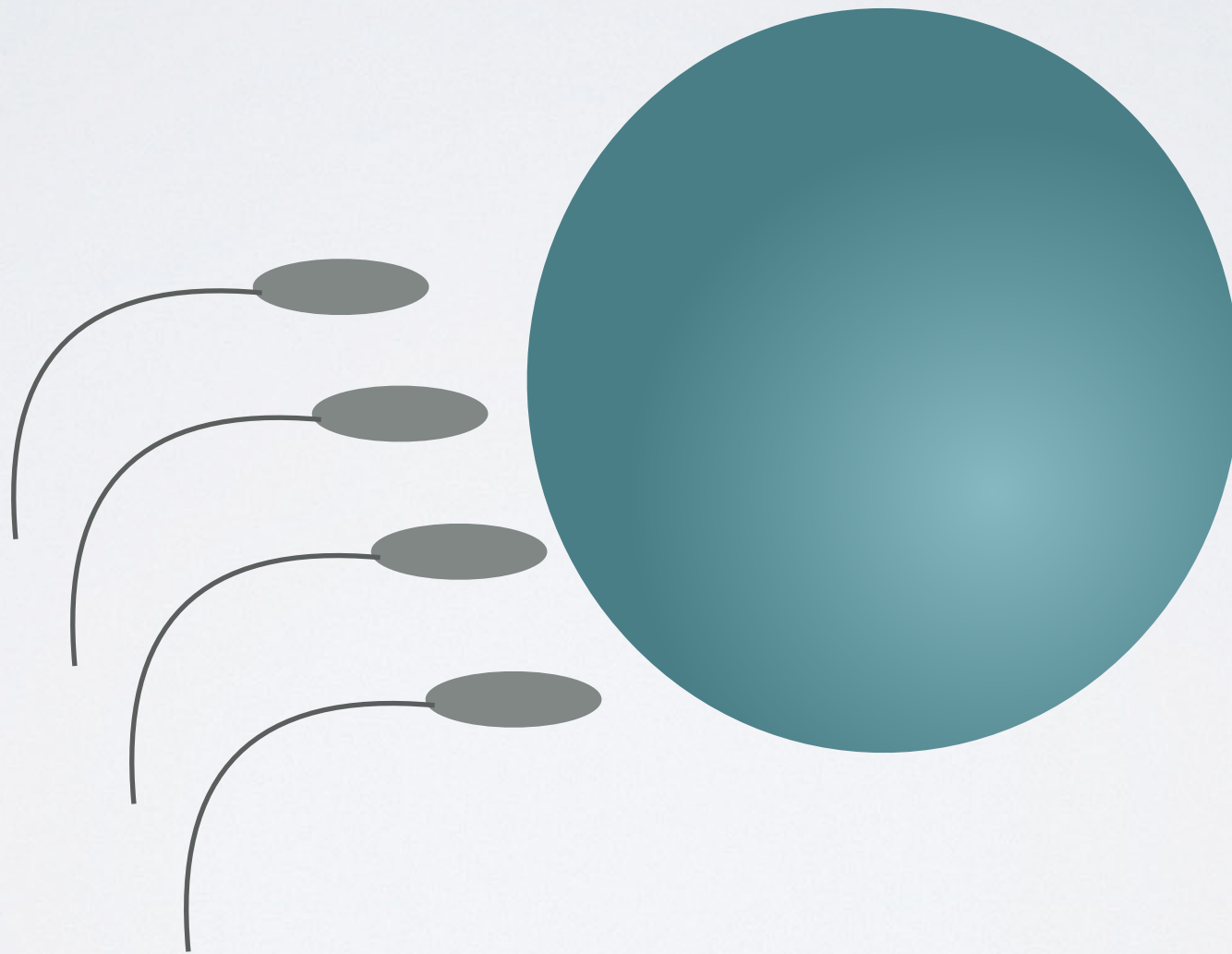
By stealing well-honed interbacterial toxins, eukaryotes get 'out-of-the-box' antibacterial immunity factors

GENOMIC VIEW OF SELECTION



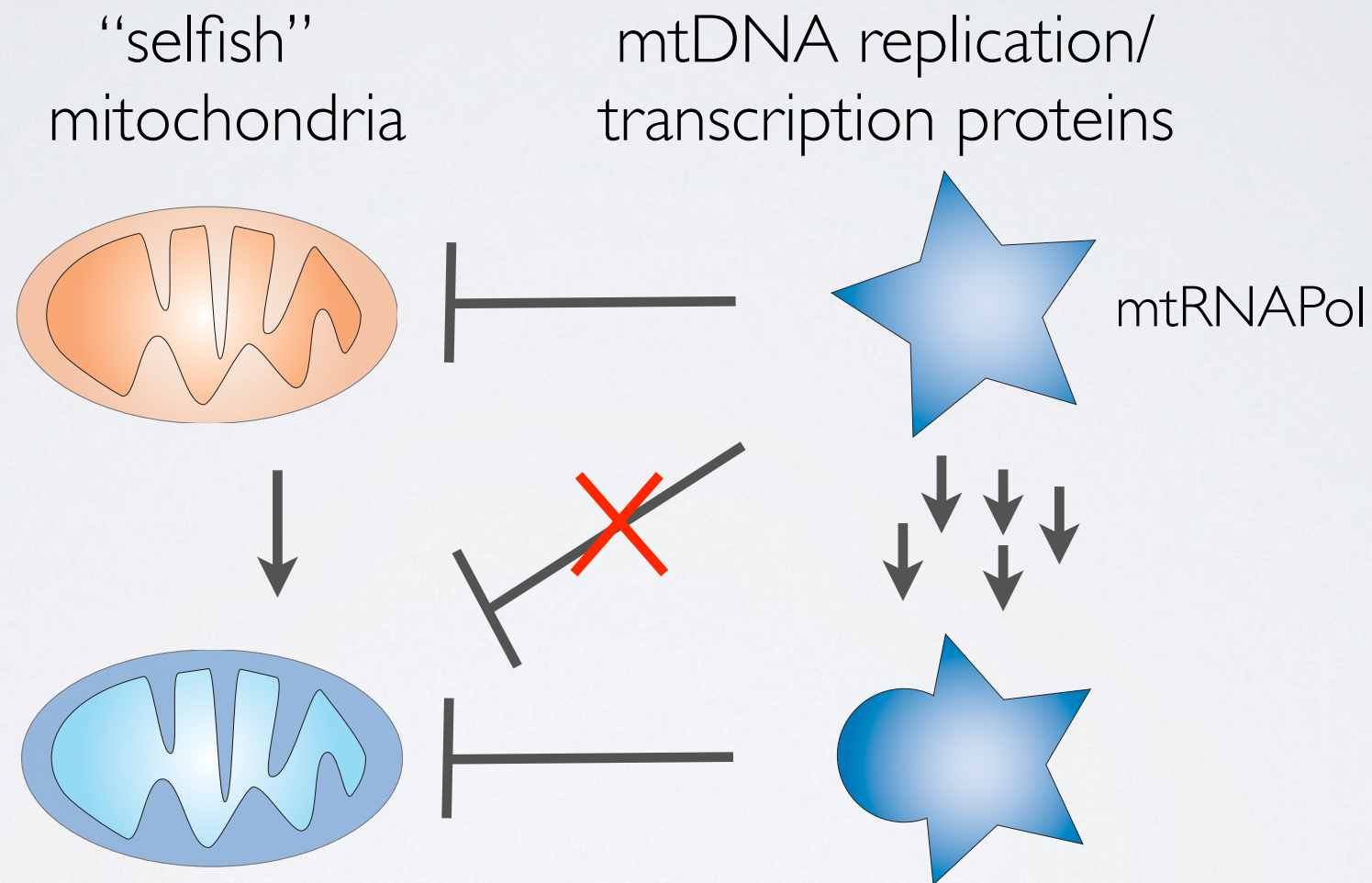
beyond the usual suspects

Genetic conflicts between components of
the same cell and even same genome



beyond the usual suspects

Genetic conflicts between components of the same cell and even same genome



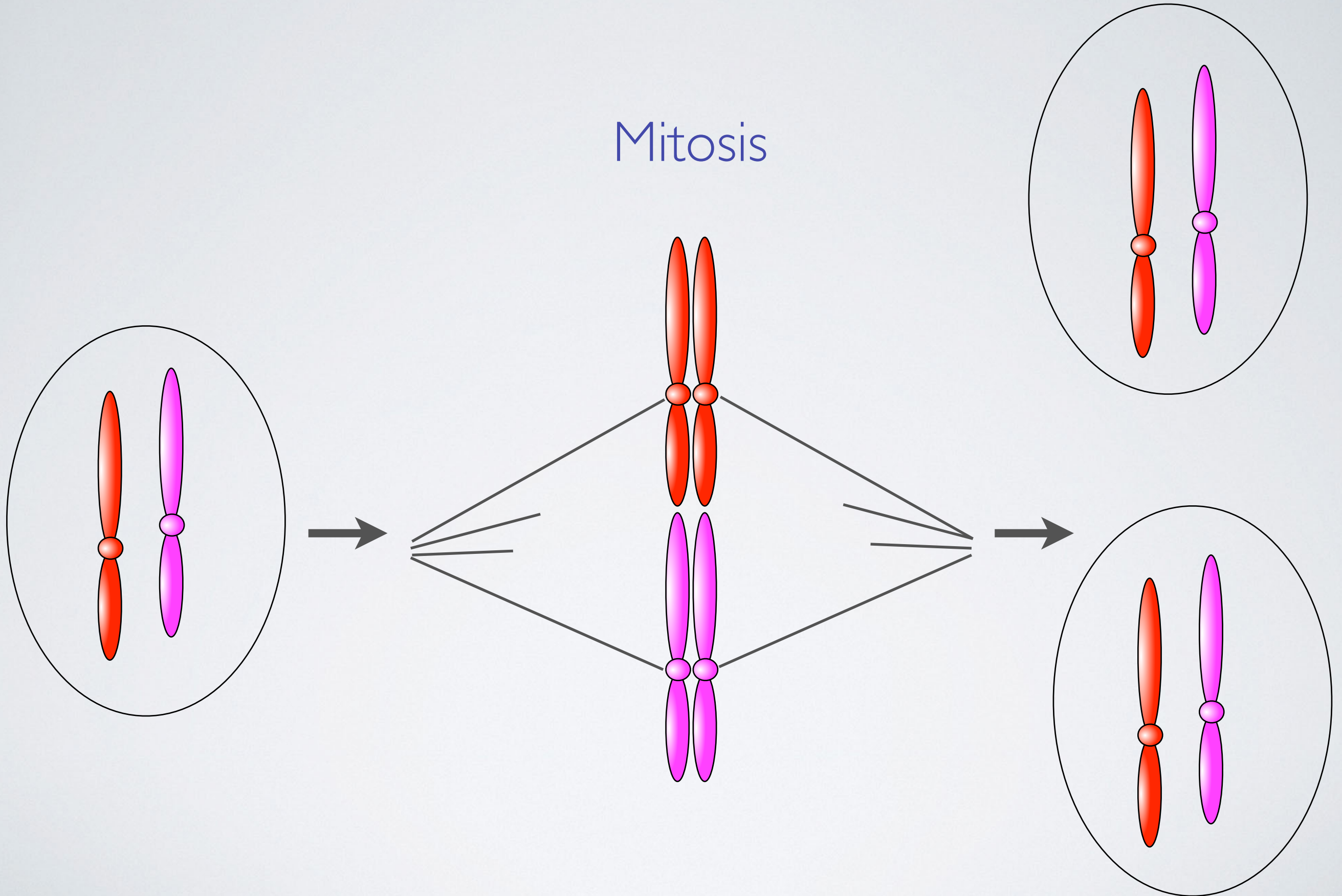
beyond the usual suspects

Genetic conflicts between components of
the same cell and even same genome



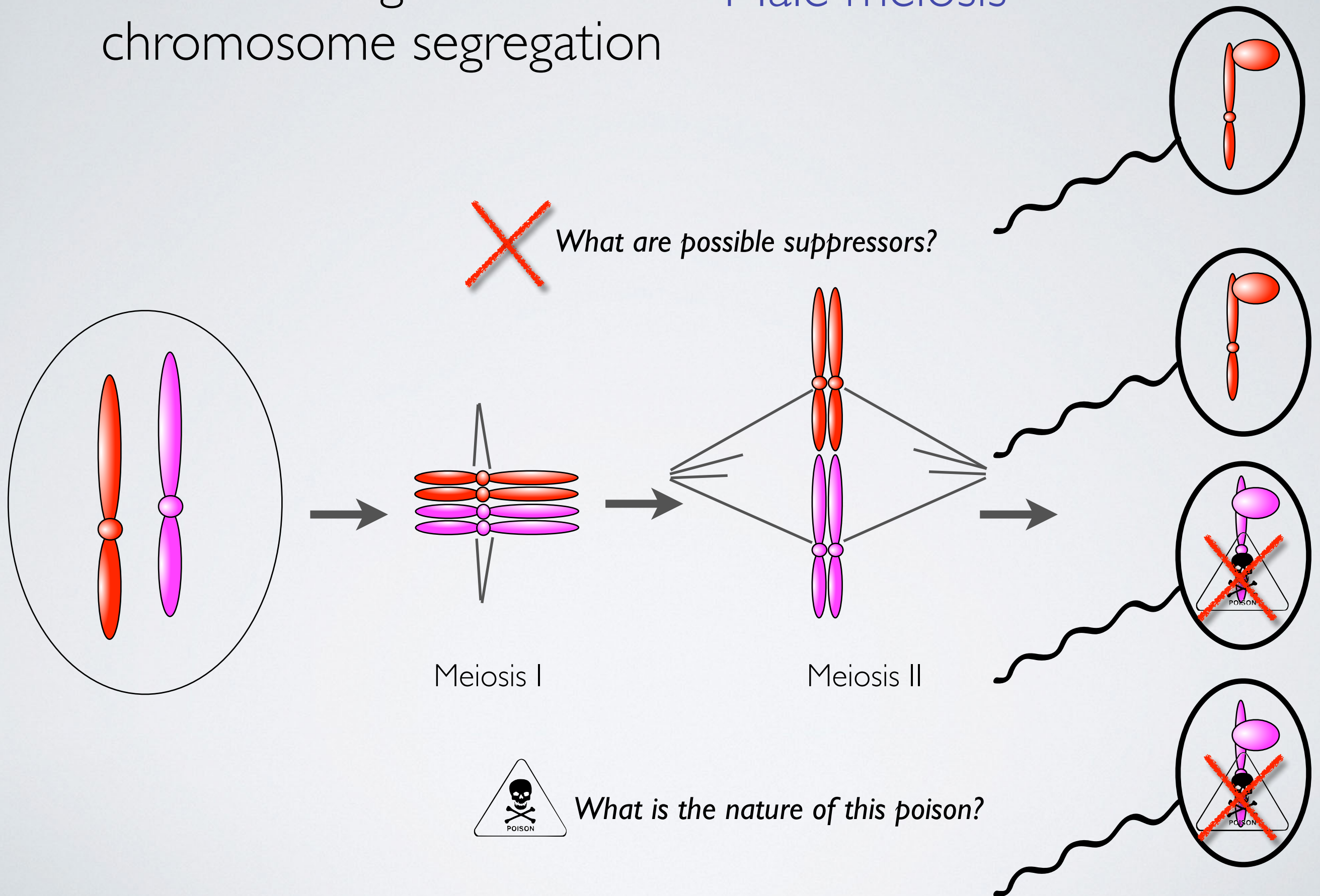
- Alleles are evolutionary competitors
 - fight fair (promote fitness; fitter allele 'wins')
 - cheat to win during/after meiosis; unfit allele can sometimes 'win'

Mitosis



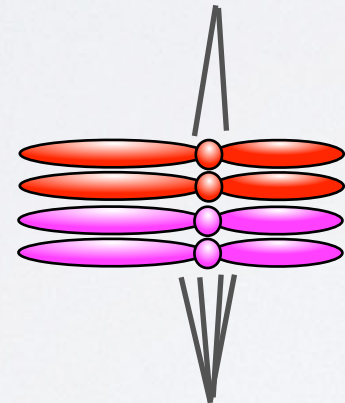
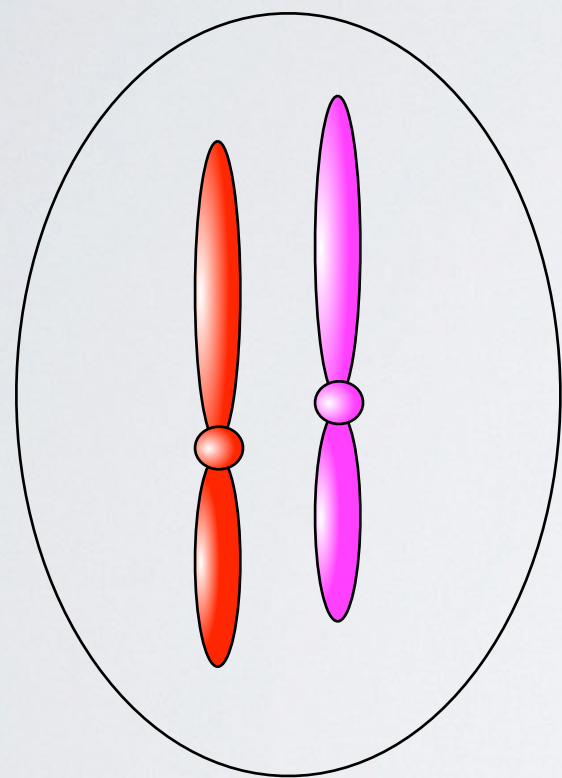
'Cheating' **after** chromosome segregation

Male meiosis

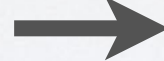


'Cheating' **during** chromosome segregation

Female
meiosis



Meiosis I



Meiosis II

oocyte
oocyte

